Companion Proceedings of the 8th International Conference on Learning Analytics & Knowledge (LAK'18)

Towards User-Centred Learning Analytics

March 5–9, 2018, Sydney, NSW, Australia

https://lak18.solaresearch.org

Organized by SOLAR SOCIETY for LEARNING ANALYTICS RESEARCH
THE UNIVERSITY OF SYDNEY

acm In-Cooperation SIGCHI sigweb

Editors:
Abelardo Pardo, The University of Sydney
Kathryn Bartimote, The University of Sydney
Grace Lynch, Society for Leanring Analytics Research
Simon Buckingham Shum, University of Technology Sydney, Australia
Rebecca Ferguson, The Open University, UK
Agathe Merceron, Beuth University of Applied Sciences, Germany
Xavier Ochoa, Escuela Superior Politécnica del Litoral, Ecuador

Cite as:
Pardo, A., Bartimote, K., Lynch, G., Buckingham Shum, S., Ferguson, R., Merceron, A., & Ochoa, X. (Eds.). (2018). Companion Proceedings of the 8th International Conference on Learning Analytics and Knowledge. Sydney, Australia: Society for Learning Analytics Research.

.

LAK 2018 Program Chairs' Welcome

It is with great pleasure that we welcome you to Sydney, Australia, for the *8th International Conference on Learning Analytics and Knowledge (LAK18).* The Conference is organised annually by the *Society for Learning Analytics Research (SoLAR)* and is hosted this year by The University of Sydney.

The aim of the conference is to provide a forum for presentation, exchange and discussion of research and practices related to the transdisciplinary field of learning analytics. This year we received 355 submissions, a record number. We offer an extensive program including a Research track (35 full papers, 26 short papers, four extended abstracts, and 17 posters and demonstrations) and a Practitioner track (11 full papers, two short papers, nine posters and demonstrations). The program begins with six full-day workshops, 13 half-day workshops and our fourth hackathon, which runs over two days. Our sixth doctoral consortium received a record-breaking 35 submissions and accepted 15 of them.

As usual, the research papers are published as an archival *Proceedings* by the ACM, but there is far more to LAK than those contributions, which is reflected in the present *Companion Proceedings*, published in open access by SoLAR. This Companion Proceedings is more comprehensive than in previous years, incorporating all workshop proceedings. In addition, here you will find practitioner papers, extended abstracts, all posters & demonstrations, the LAK hackathon and doctoral consortium papers.

We are delighted to welcome our Keynote Speakers, chosen to stretch our thinking in new directions: Prof. David Williamson Shaffer on quantitative ethnography, Prof. Cristina Conati on personalising visualisation, and Prof. Neil Selwyn on questioning our blind spots as a community.

The theme for LAK18 is *Towards User-Centred Analytics.* An important feature of the LAK community that attracts diverse delegates is our interest in the human factors in learning analytics systems. As learning analytics tools move out of the lab into the real world, their success or failure must be judged not only on technical criteria, but also by their adoption and effectiveness in schools, universities and workplaces. Often this is where the gulf between hype and reality becomes apparent. The complexities of embedding innovative technology in authentic contexts open a range of critical challenges for the field.

While LAK has always encouraged contributions dealing with issues related to adoption, LAK18 places particular emphasis on how various stakeholders can, or must, be engaged in the design, deployment and assessment of learning analytics tools and policies, if they are to be successful and sustainable. In order to do this, we have invited theoretical, methodological, empirical and technical contributions addressing topics including:

- Which design processes involve learners, educators and other users effectively in the co-design of analytics tools?
- Which techniques are effective in assessing how end-users make sense of, interact with, and act on analytics feedback?
- In what ways can learning analytics systems be biased, and can they be made more transparent and accountable to different stakeholder groups?
- How are educational leaders creating the conditions for learning analytics systems to take root and grow?
- How strong is the evidence that the adoption of learning analytics benefits stakeholders?

This theme is reflected in the program with two sessions on User-Centred Design.

As a conference grows, especially one as multidisciplinary as LAK, it is a challenge to maintain the scientific quality. We put significant effort into matching submissions with reviewer expertise, and ensured that all papers received three reviews. We have also introduced two innovations in the review process. First, the evidence is that double-blind reviewing increases reviewer objectivity, so this was the first LAK to disguise authorship. Second, the Program Committee all acted as meta-reviewers, overseeing initial reviews and encouraging reviewer discussion in order to resolve differences, before making recommendations to the Program Chairs. While no peer-review process is perfect, our view is that both of these measures were effective in increasing reviewer objectivity, and in scaling the peer-review process.

There are several other innovations this year. The first is the addition of the *Extended Abstract* as a submission category. This was designed to encourage submissions from researchers previously under-represented at LAK, who can bring new perspectives. The second is the introduction of this full *Companion Proceedings*, archived on the SoLAR website, incorporating practitioner papers, extended abstracts, workshop outlines and papers, descriptions of demonstrations, and poster abstracts. Thirdly, recognising the relative absence of analytics work in primary and secondary schools at LAK, we are very pleased that this year sees the first *Analytics in Schools* workshop, helping to build this community. Finally, an invitational *Leadership Track* is running on Wednesday, starting with a plenary panel, to build the network of senior institutional leaders whose role in enabling learning analytics is so critical.

Finally, we are indebted to the 83 members of the Program Committee and the 251 Reviewers for their thoughtful and helpful reviews, comments and meta-reviews. Their work was not easy given the diversity and high quality of the works under review. Only with their support are we able to provide you with this program for LAK18.

Simon Buckingham Shum, University of Technology Sydney, Australia
Rebecca Ferguson, The Open University, UK
Agathe Merceron, Beuth University of Applied Sciences, Germany
Xavier Ochoa, Escuela Superior Politécnica del Litoral, Ecuador (ex-officio)

LAK 2018 Conference Organization

Conference Chairs

Abelardo Pardo, The University of Sydney, AUS
Kathryn Bartimote-Aufflick, The University of Sydney, AUS
Grace Lynch, Society for Learning Analytics Research, AUS

Program Chairs

Simon Buckingham Shum, University of Technology Sydney, AUS
Rebecca Ferguson, The Open University, UK
Agathe Merceron, Beuth University of Applied Sciences, GER
Xavier Ochoa, Escuela Superior Politécnica del Litoral, EC (ex-officio)

Practitioner Track Chairs

John Whitmer, Blackboard Inc., USA
Cassandra Colvin, Charles Sturt University, AUS

Workshop Chairs

Linda Corrin, The University of Melbourne, AUS
Niels Pinkwart, Humboldt-Universität Berlin, GER

Doctoral Consortium Chairs

Ani Aghababyan, McGraw-Hill Education, USA
Judy Kay, University of Sydney, AUS
Lori Lockyer, University of Technology Sydney, AUS
Bart Rienties, The Open University, UK
Kalina Yacef, University of Sydney, AUS

Posters and Demonstrations Chairs

Catherine Spann, University of Texas at Arlington, USA
Michael Eagle, Carnegie Mellon University, USA

Local Chair

Brooke Fuz, The University of Sydney, AUS

Program Committee for Companion Proceedings

Adam Cooper	Tribal	U.K.
Bill Searle	Charles Darwin University	Australia
Britte Cheng	SRI International	U.S.A.
Cassandra Colvin	Charles Sturt University	Australia
Catherine Gunn	The University of Auckland	New Zealand
Danny Y.T. Liu	The University of Sydney	Australia
David Jones	University of Southern Queensland	Australia
Deb West	Flinders University	Australia
Didem Tufan	Middle East Technical University	Turkey
Hazel Jones	University of Southern Queensland	Australia
James Williamson	University of California, Los Angeles	U.S.A.
John Fritz	University of Maryland, Baltimore County	U.S.A.
John Whitmer	Blackboard, Inc.	U.S.A.
Kimberly Arnold	University of Wisconsin-Madison	U.S.A.
Lalitha Agnihotri	McGraw Hill Education	U.S.A.
Lin Zhou	IBM	U.S.A.
Maria Ofelia San Pedro	ACT, Inc.	U.S.A.
Marie Bienkowski	SRI International	U.S.A.
Marion Blumenstein	The University of Auckland	New Zealand
Mary Jean Blink	TutorGen, Inc.	U.S.A.
Melinda Lewis	Charles Sturt University	Australia
Mengxiao Zhu	Educational Testing Service	U.S.A.
Niall Sclater	Sclater Digital	U.K.
Nicholas Lewkow	McGraw-Hill Education	U.S.A.
Patsy Moskal	University of Central Florida	U.S.A.
Sakinah Alhadad	Griffith University	Australia
Serge Garlatti	IMT Atlantique	France
Shelley Kinash	University of Southern Queensland	Australia
Shirin Mojarad	McGraw Hill Education	U.S.A.

Stephen Fancsali	Carnegie Learning, Inc.	U.S.A.
Timothy Harfield	Blackboard Inc	U.S.A.
Xavier Ochoa	Escuela Superior Politécnica del Litoral	Ecuador
Yong-Sang Cho	Korea Education & Research Information Service	South Korea

Posters and Demonstrations

Adrian Holzer	Ecole Polytechnique Fédérale de Lausanne	Switzerland
Agnes Sandor	Naver Labs Europe	France
Albrecht Fortenbacher	HTW Berlin	Germany
Ana Fernández Vilas	University of Vigo	Spain
Aneesha Bakharia	The University of Queensland	Australia
Barry Fishman	University of Michigan	U.S.A.
Behnam Taraghi	Graz University of Technology	Austria
Caitlin Mills	The University of British Columbia	Canada
Caroline Haythornthwaite	Syracuse University	U.S.A.
Carolyn Rose	Carnegie Mellon University	U.S.A.
Catherine Spann	University of Texas at Arlington	U.S.A.
Cecilia Aguerrebere	Duke University	U.S.A.
Cristian Cechinel	UFSC – Federal University of Santa Catarina	Brazil
Cristóbal Cobo	Centro de Estudios Fundación Ceibal	Uruguay
Cristobal Romero	University of Cordoba	Spain
Daniel Spikol	Malmö University	Sweden
Demetrios Sampson	Curtin University	Australia
Dirk Ifenthaler	University of Mannheim	Germany
Drew Paulin	University of California, Berkeley	U.S.A.
Elle Yuan Wang	Columbia University	U.S.A.
Enayat Rajabi	Dalhousie University	Canada
François Bouchet	LIP6 – Université Pierre et Marie Curie	France
Fridolin Wild	Oxford Brookes University	U.K.
Gábor Kismihók	University of Amsterdam	Netherlands
Gregor Kennedy	The University of Melbourne	Australia
Hassan Khosravi	The University of Queensland	Australia
Hendrik Drachsler	The Open University	Netherlands

Hiroaki Ogata	Kyoto University	Japan
Hiroyuki Yamada	Carnegie Foundation for the Advancement of Teaching	U.S.A.
Huzefa Rangwala	George Mason University	U.S.A.
Ido Roll	The University of British Columbia	Canada
Irene-Angelica Chounta	Carnegie Mellon University	U.S.A.
Jakub Kužílek	CTU in Prague	Czech Republic
Justin Reich	Massachusetts Institute of Technology	U.S.A.
Kairit Tammets	Tallinna Ülikool	Estonia
Karina Cela	Escuela Politécnica del Ejército	Spain
Kenneth Holstein	Carnegie Mellon University	U.S.A.
Kevin Mayles	The Open University	U.K.
Laura Allen	Arizona State University	U.S.A.
Mahdi Bashari	University of New Brunswick	Canada
Maren Scheffel	Open Universiteit	Netherlands
Mathieu D'Aquin	National University of Ireland Galway	Ireland
Matt Demonbrun	University of Michigan	U.S.A.
Michael Eagle	Carnegie Mellon University	U.S.A.
Michael Yeomans	Harvard University	U.S.A.
Michel Desmarais	Ecole Polytechnique de Montreal	Canada
Mihai Dascalu	University Politehnica of Bucharest	Romania
Mike Tissenbaum	Massachusetts Institute of Technology	U.S.A.
Mykola Pechenizkiy	Eindhoven University of Technology	Netherlands
Negin Mirriahi	University of South Australia	Australia
Nia Dowell	The University of Memphis	U.S.A.
Ouajdi Manai	The Carnegie Foundation for the Advancement of Teaching	U.S.A.
Patrick Jermann	Ecole Polytechnique Fédérale de Lausanne	Switzerland
Philippe Dessus	LSE, Univ. Grenoble Alpes	France
Rebeca P. Díaz-Redondo	University of Vigo	Spain
Ricardo Matsumura Araujo	Federal University of Pelotas	Brazil
Ryan Baker	University of Pennsylvania	U.S.A.
Sanam Shirazi Beheshtiha	The University of British Columbia	Canada
Sandra Milligan	Univerity of Melbourne	Australia
Sandro Camargo	UNIPAMPA – Universidade Federal do Pampa	Brazil

Seth Adjei	Worcester Polytechnic Institute	U.S.A.
Sharon Slade	The Open University	U.K.
Stefan Trausan-Matu	University Politehnica of Bucharest	Romania
Stian Håklev	Ecole Polytechnique Fédérale de Lausanne	Canada
Sven Charleer	Katholieke Universiteit Leuven	Belgium
Tanja Käser	Stanford University	U.S.A.
Thomas Daniel Ullmann	The Open University	U.K.
Tobias Hecking	University of Duisburg-Essen	Germany
Tobias Ley	Tallinn University	Estonia
Ulla Lunde Ringtved	UCN, University College of Northern Denmark	Denmark
Vitomir Kovanovic	The University of Edinburgh	U.K.
Yeonjeong Park	Honam University	South Korea
Yi-Shan Tsai	The University of Edinburgh	U.K.

Table of Contents

(click to go to paper)

Extended Abstracts

Practitioner Papers

Practitioner Posters

Doctoral Consortium

Posters

Demonstrations

Learning Analytics in Schools

Workshops

Extended Abstracts

Considering Context and Comparing Methodological Approaches in Implementing Learning Analytics at the University of Victoria

Author(s): Sarah K. Davis, Rebecca L. Edwards, Mariel Miller, and Janni Aragon

ABSTRACT: One of the gaps in the field of learning analytics is the lack of clarity about how the move is made from researching the data to optimizing learning (Ferguson & Clow, 2017). Thus, this report details the implementation process undertaken between the data to the metrics of the learning analytics cycle (Clow, 2012). Five anonymized secondary data sets consisting solely of LMS interaction data from undergraduate courses at a large research university in Canada university will be analyzed. Specifically, this study (a) provides context for the individual data sets through a survey tool taken by the instructors of the course, and (b) compares machine learning techniques and statistical analyses to provide information on how different approaches to analyzing the data can inform the learning process. Findings from this study will inform the adoption of learning analytics at the institution and contribute to the learning analytics community by detailing the methods compared in this report.

Keywords: learning analytics, higher education, implementation, context, methods

1 INTRODUCTION

Implementing learning analytics at any institution is not a one-size-fits-all approach (Rienties, 2017, August). Rarely can learning analytics systems be transferred directly from one institution to another due to a variety of factors. However, Ferguson and Clow (2017) posit one of the gaps in the field of learning analytics is the lack of clarity about how the move is made from researching the data to optimizing learning in implementation. This practioner's report aims to fill this gap by detailing the steps we took during this part of our project where we were researching how (a) to use the data collected from the LMS during learning, and (b) to evaluate what approaches would best support learning. Specifically, this study details the research undertaken between the data to the metrics of the learning analytics cycle (Clow, 2012).

This paper is a part of a larger Learning Analytics project that began in 2016 when a specific academic service department at the university initiated exploring the use of learning analytics. This institution is located in Canada and has a student enrollment of 20 000 undergraduate and graduate students. At our institution, our current learning analytics project has three phases. Phase 1, completed in 2016-2017, was to complete a literature review of learning analytics in Canadian universities, as well as a general overview of what learning analytics could offer to the institution. This paper focuses on Phase 2 of the project, which is the evaluation of how learning analytics data can be analyzed, used, and organized (see figure 1). Phase 3, which relies on the results of Phase 2, will run a pilot of learning analytics in 2018-2019 along with a visualization dashboard to selected undergraduate courses. The aims of the current phase are to (a) provide context for the LMS data for five undergraduate courses through a survey tool taken by the course instructors, and (b) compare different methodological approaches and analyses based on previous research.

Figure 1: Illustration of placement of current phase of project

2 PERSPECTIVE / APPROACH / METHODOLOGY

One of the many challenges in using learning analytics is incorporating a theoretical framework in order to move "from clicks to constructs" (Knight & Buckingham Shum, 2017, p. 17). Considering the epistemology, pedagogy, and assessment of the types of learning analytics used, the purposes, and the audience for the data will help articulate and address the types of knowledge targeted within the LMS, how that knowledge will be used, and by whom. For this project, we considered all three categories and how our project will address them (see table 1).

Table 1: Application of Knight & Buckingham Shum's (2017) Theoretical Framework

Theory	Design Questions	Our Learning Analytics Project
Epistemology	What are we measuring?	We are measuring how students are engaging within the LMS on various course components.
	How are we measuring?	We are measuring engagement through the amount of times students view/complete various course components in the LMS and how these views/completions compare with other students.
Pedagogy	Why is this knowledge important to us?	To both identify and support students at risk of failure and to promote success among students.
	For whom is the analytic?	The analytics are for students, instructors, and advisors to inform learning, teaching, and advising practices.
Assessment	Where does the assessment/feedback happen?	The assessment/feedback happens as students are using the LMS to complete course components both as required on the syllabus and above and beyond what is required.

Our other theoretical framework, student engagement, originates from educational psychology and attempts to articulate students' feelings, beliefs, thoughts, and behaviours in academic environments. There are 4 factors of student engagement: behavioural, emotional, cognitive, and agentic (Fredricks, Blumenfeld, & Paris, 2004; Reeve, 2013; Sinatra, Heddy, & Lombardi, 2015). Davis, Edwards, Hadwin, & Milford (2017), in their study examining how log file data could distinguish between high and low performers in a large undergraduate elective course, found differences across student engagement, particularly behavioural and agentic engagement. Their results replicated Reeve's (2013) findings that behavior and agentic engagement are predictive of achievement, however with log file data, not self-reports of engagement as used in the original study. Translated to learning analytics, behavioural engagement denotes LMS log file activity required by the syllabus, such as attendance, quizzes, etc, and agentic engagement is defined as actions in the LMS not connected to graded activities, such as days viewed course, syllabus views, etc (Davis et al., 2017; Edwards, Davis, Hadwin, and Milford, 2017;)

We had three questions guiding this phase of the implementation project:

1. What data within the LMS are meaningful indicators of student success in a course?
2. How do responses from instructors on a contextual survey inform how the data is organized for analysis?
3. How do three iterations of methodological approaches to learning analytics data compare in providing indicators of student performance in large, first year undergraduate courses?

3 DESIGN / IMPLEMENTATION

As of February 2018, we have received ethical approval to use the secondary data from the LMS, and recruited three instructors who have given us permission to use the data from their courses from the previous academic year. We have administered the instructor context survey and prepared all the data for analysis. The next steps will be to compare the methodological approaches using machine learning (i.e. Naïve Bayes) and regression analyses to determine indicators of student success in our data set. The main questions from the survey are below and were informed by previous research (Edwards et al., 2017; Davis et al., 2017):

1. How important is the use of your Moodle site to succeeding in your course?
2. How important were each of the following Moodle components in your course?
3. How many days per week on average did you expect students in your course to access the Moodle site?
4. What activities embedded in Moodle were required components of your course (i.e. in your syllabus)? What activities were optional (i.e. not in your syllabus)?
5. What other information about how you used your Moodle site would be helpful for us to know?
6. Briefly describe how you provided grades to students.
7. What essential components of your course were not part of this Moodle course shell?

4 IMPLICATIONS / FUTURE WORK

Our report aims to reveal the processes one institution is using in its investigation into learning analytics through revealling (a) how different analyses can highlight indicators of student performance from LMS data, and (b) how a survey filled out by instructors can provide contextual parametres to LMS data. We are in the process of conducting the analyses and therefore will present our findings at LAK 2018.

The results of the instructor survey will provide vital contextual information about the data from the LMS. Limited previous research has relied on instructor reports of the varying importance of LMS data, the requirements of the course regarding student activity as detailed in the syllabus, or the number of days students would need to access the LMS to complete course requirements. By examining how this survey data influences the organization and therefore the interpretation of data will inform the next phase of this project and ultimately contribute to learning analytics research that seeks to improve student performance, and for other practioners interested in different approaches.

Previous research highlights what institutions have found when analyzing data, but there is a lack of research on how this move was made or why certain methods were chosen over others. Comparison of methodological approaches is beneficial when moving from data to implementation; we looked at both how different educational theories and analyses can affect implementation decisions. Our project has the potential to inform other practioners about the challenges in the early stages of implementation, including ethics, approaches, and analyses, all of which may vary depending on the goals of the institution in its use and adoption of learning analytics.

REFERENCES

Clow, Doug (2012). The learning analytics cycle: closing the loop effectively. *Proceedings of the Second International Learning Analytics & Knowledge Conference*, pp. 134-138. New York: ACM. DOI: 10.1145/2330601.2330636

Davis, S.K., Edwards, R.L., Hadwin, A.F., & Milford, T.M. (2017, May). An exploratory study of a trimodal distribution of student achievement in an undergraduate self-regulated learning course. Paper presented at annual meeting of the Canadian Society for the Study of Education Conference (CSSE), Toronto, Ontario.

Edwards, R. L., Davis, S.K., Hadwin, A.F., & Milford, T.M. (2017). Using predictive analytics in a self-regulated learning university course to promote student success. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, pp. 556-557. New York: ACM. DOI: 10.1145/3027385.3029455

Ferguson, R. & Clow, D. (2017). Where is the evidence? A call to action for learning analytics. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, pp. 56-65. New York: ACM. DOI: 10.1145/3027385.3027396

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research,* 74(1), 59-109. DOI: 10.3102/00346543074001059

Knight, S. & Buckingham Shum, S. (2017). Theory and learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.) *Handbook of Learning Analytics*, 1st ed. SOLAR. DOI: 10.18608/hla17.001

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology,* 105(3), 579-595. DOI: 10.1037/a0032690

Rienties, B. (2017, August). *The power of learning analytics: Is there still a need for educational research?* Keynote presented at the European Association for Research on Learning and Instruction Conference, Tampere, Finland.

Sinatra, G.M., Heddy, B.C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, 50(1), 1-13. DOI: 10.1080/00461520.2014.1002924

Using an iterative, multi-stakeholder process to develop an experience-driven, institution-wide learning analytics policy

Jan-Paul van Staalduinen
Delft University of Technology, the Netherlands
j.p.vanstaalduinen@tudelft.nl

ABSTRACT: Implementing learning analytics in an institution without a well-developed and appropriate policy and support from the many internal stakeholders brings inherent risks and challenges. This paper describes the iterative, multi-stakeholder approach that was used by Delft University of Technology (Netherlands) to create an institution-wide learning analytics vision and policy document, and the findings, outcomes, and lessons learned from that process. The chosen iterative approach for learning analytics policy development consisted of four phases: 1) make an inventory of external insights and lessons learned; 2) determine priorities of the organisation; 3) develop policy principles; and 4) determine recommended future steps. During each of these phases both external experts and internal stakeholders were consulted. The learning analytics policy development process resulted in a number of findings and outcomes from each of the four phases of the process, the most important of which were a policy document and a series of next steps for undertaking a campus-wide implementation of learning analytics in education.

Keywords: Learning Analytics Policy; Values, Ethics and Law; Adoption.

1 INTRODUCTION

The field of learning analytics is developing and evolving at a rapid pace (Bienkowski, Feng, and Means, 2012; European Commission, 2014; ECAR-Analytics Working Group, 2015; European Commission, 2015; Higher Education Commission, 2016). The use of learning analytics offers opportunities, but also risks for a university (Buckingham Shum, 2012; Slade and Prinsloo, 2013). For the well-considered application of learning analytics in campus education, not only should privacy and ethical issues be addressed, but clear guidelines are needed to reassure both students and academic staff, and to provide a clear ruleset for the deployment of learning analytics (Elouazizi, 2014; Open University, 2014a; Pardo and Siemens, 2014; Sclater and Bailey, 2015; Khalil and Ebner, 2015). Implementing learning analytics in an institution without a well-developed and appropriate policy brings inherent risks and challenges (Higher Education Commission, 2016).

After a few years of experimenting with small-scale learning analytics pilots in online education, the executive board of Delft University of Technology (Netherlands) has decided to make the introduction of learning analytics in campus education a major focus for the academic year 2017-2018. The university has clear ambitions for learning analytics, as part of a broader focus on evidence-based education, and there is a strong interest within the organisation for the subject.

To provide a supportive framework for future on-campus learning analytics projects, which are to be piloted and deployed in the coming years, a policy for learning analytics needed to be created. As learning analytics touches upon students' personal lives, this creates a need for an easy to find, transparent policy, to provide clarity during the university's explorative phase into a valid use of learning analytics methods and interventions – and beyond.

When it comes to developing learning analytics policy, Campbell and Oblinger (2007) stress the importance of involving all stakeholders as the introduction of learning analytics will not only affect students, but also campus staff. To accommodate for this, in line with Greller and Drachsler (2012), the university adopted an iterative, multi-stakeholder approach for learning analytics policy development, where external experts and literature were consulted for best practices in learning analytics policy, whilst simultaneously discussions and conversations were held with internal stakeholders. A key aspect of this process was the university's desire to speed up its internal learning process by making use of experiences and lessons learned from universities that were ahead in this field (e.g. Tsai and Gasevic, 2017). This paper describes the approach taken and its outcomes.

2 ITERATIVE, MULTI-STAKEHOLDER POLICY DEVELOPMENT

The university's chosen approach for learning analytics policy development consisted of four phases: 1) make an inventory of external insights and lessons learned; 2) determine priorities of the organisation; 3) develop policy principles; and 4) determine recommended future steps. The policy development process started in December 2016 and was concluded in October 2017.

To create an overview of what other institutions were doing in this area, available learning analytics policies of other institutes were carefully analysed (Open University, 2014a; Open University, 2014b; Sclater and Bailey, 2015; University of Edinburgh, 2017; University of West London, 2017). In addition, external experts in the field were interviewed.

The university's organisational priorities were determined through a series of discussions that were held with key internal stakeholders, from university leadership, university support staff, the university's ethics committee, the Student Council (elected representatives of the student population), and academic staff. In addition, during each of the phases of the policy development process iterating versions of the policy document were discussed with these internal stakeholders, in order to incorporate their feedback and accommodate for their interests in the final version.

The output of the development approach was a policy document which contained the university's vision for learning analytics on campus, guidelines and rulesets that applied to the implementation of learning analytics in campus education, and a series of next steps to gradually evolve the university's capacity for and knowledge of learning analytics. This policy document would contain clear guidelines and rulesets, but at the same time would offer sufficient freedom for experimentation and organisational learning.

3 FINDINGS AND OUTCOMES OF THE POLICY DEVELOPMENT PROCESS

The learning analytics policy development process resulted in a number of findings and outcomes from each of the four phases.

3.1 Inventory of external insights and lessons learned

From the literature a general consensus could be distilled about the topics that should be covered by learning analytics policies, namely: ethics and privacy (also in a legal context); data protection; data access, usage and purpose of usage; transparency about how data is used; acquiring student consent; and communicating with stakeholders (Elouazizi, 2014; Pardo and Siemens, 2014; Sclater and Bailey, 2015; Khalil and Ebner, 2015, Higher Education Commission, 2016).

The interviewed experts also advocated the use of a clear and focused strategy for implementing learning analytics. This strategy should not only cover how learning analytics will be used in campus education, but moreover should strongly emphasize how students, teaching staff, and other organisational players are involved during implementation, in order to allow internal stakeholders to build a shared awareness and knowledge of learning analytics, and to create internal support.

3.2 Priorities of the organisation

Interviews with internal stakeholders found a high level of support for the use of learning analytics in campus education, with executives, faculty staff, and the student population strongly displaying interest and enthusiasm. Students saw and embraced the potential benefits of learning analytics, although they did have concerns about ethics and privacy. The prevailing sentiment among the students was that learning analytics systems should only be used to provide support and feedback to students, but should not for example be used to put additional study demands on students. There was also an emphasis on transparency on data collection and use, as the expectation was transparency would take away many of the concerns students had. This meant that the university should be open about which data is collected about students; how it is used in analyses; and what form of data use an individual student has consented to.

In addition, faculty staff voiced concerns about whether the use of learning analytics would be all-inclusive: they expected learning analytics to be useful for many students, but argued that there would always be a group of students who would be out of reach of interventions and thus could not be supported this way. Faculty staff also opined that the university had a good foundation to build on, through experiences from its own pilot activities in this area, cooperation with other universities in this area, and the existing expertise in supporting and academic staff.

3.3 Learning analytics policy principles

In the final version of the policy document, fourteen policy principles were proposed for the use of learning analytics in campus education. These emphasize learning analytics as a discipline with ethical dimensions, the application of which must be in line with the core values of the university. They also state that the university should use learning analytics to provide targeted support to all students, not just at-risk ones. In addition the principles state that it is the university's duty to be transparent about how data is collected, stored, and used for analysis, including used methods and algorithms in the any applications of learning analytics that are used in campus education.

Additional principles focus on the necessity to explicitly request students' consent to use their personal data for analytics purposes. They also regulate the continual involvement of teaching staff and students in developing and evaluating the use of learning analytics and their governance.

3.4 Recommended future steps

To determine the university's next steps, from the insights gained through the policy development process, five points were formulated for the campus-wide roll-out of learning analytics in education:

1. Create a strategic project plan: The university should determine at the institutional level how to introduce learning analytics on-campus, and develop an implementation strategy for this. This would require making conscious and specific choices, for which a multiannual project plan would provide a structured approach.

2. Create an advisory board for learning analytics: A permanent advisory committee with experts, teaching staff, and student representatives should be created for maintaining support. This committee should serve as an advisory body for any university learning analytics projects.

3. Make work of transparency and consent: In order to give institutional transparency about consent and data protection a prominent place, an informative website should be created. At the same time, the university should put effort into making it technically possible for students to actively manage their consent settings in the university's administrative systems.

4. Start with small projects to gain experience: As learning analytics is a still-evolving discipline, there is an inherent risk in launching large-scale projects. Starting with a number of small pilots and experiments focused on simple statistics and expanding them slowly allows for the university to gain experience, and to discover which interventions work, what data is needed, and how to align initiatives within existing legal and policy frameworks.

5. Increase knowledge and awareness about learning analytics within the organization: The university should organize regular meetings with students and staff, and use these meetings for formative and educational purposes, so that an informed debate within the university about the use of learning analytics on education can be held.

4 IMPLICATIONS AND FUTURE WORK

The iterative, multi-stakeholder policy development process resulted in an institutional policy that was approved by the executive board. In addition, a number of key learnings came from the policy development process.

A first lesson learned was the importance of aligning experimental learning analytics projects with existing legal and policy frameworks. This provides a tension and requires a balancing act to do right.

A second lesson learned was that the university's student population was generally curious, interested, and enthusiastic about the introduction of learning analytics on-campus and its potential benefits, yet at the same time had concerns about certain risks, primarily with regards to privacy. Keeping them informed and involved, and addressing their concerns in formal policy and through the use of the advisory board, should greatly increase the chances of a successful implementation.

A third lesson learned was the realisation that learning analytics touches upon many legal and ethical frameworks, but that legal professionals with expertise in all those fields are rare to find, meaning there is a knowledge gap to be covered in that domain.

A campus-wide roll-out of learning analytics in education is expected to be a challenging project, as it requires alignment of the university's many internal departments and faculties, and managing data from the complex ecosystem of administrative and educational systems. This will require years of dedicated effort, while expectations are high.

5 ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the Erasmus+ Programme of the European Union; the Erasmus+ Forward Looking Cooperation STELA Project, number 562167-EPP-1-2015-BE-EPPKA3-PI-FORWARD.

REFERENCES

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology,* 1, 1-57.

Buckingham Shum, S. (2012). Learning analytics [policy brief]. Moscow, RU: United Nations Educational, Scientific and Cultural Organization UNESCO Institute for Information Technologies in Education.

Campbell, J. P., & Oblinger, D. G. (2007). Academic analytics. *EDUCAUSE review, 42*(4), 40-57.

Elouazizi, N. (2014). Critical Factors in Data Governance for Learning Analytics. *Journal of Learning Analytics, 1*(3), 211-222.

ECAR-Analytics Working Group. (2015). The predictive learning analytics revolution: Leveraging learning data for student success. *ECAR working group.*

European Commission (2014). Report to the European Commission on New modes of learning and teaching in higher education. October 2014. *Luxembourg: Publications Office of the European Union. ISBN 978-92-79-39789-9*

European Commission (2015). Education and Training Monitor 2015. *Luxembourg: Publications Office of the European Union. ISBN 978-92-79-50620-8*

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society, 15*(3), 42.

Higher Education Commission (2016). From Bricks to Clicks - The Potential of Data and Analytics in Higher Education. *Policy Connect, London.*

Khalil, M., & Ebner, M. (2015). Learning analytics: principles and constraints. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (pp. 1326-1336)*.

Open University (2014a). Ethical use of Student Data for Learning Analytics Policy FAQs *(Online only). Retrieved, March 1, 2016, from: http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-student-data-faq.pdf*

Open University (2014b). Policy on Ethical use of Student Data for Learning Analytics *(Online only). Retrieved, March 1, 2016, from: http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-use-of-student-data-policy.pdf*

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438-450.

Slade, S. and Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. *American Behavioral Scientist, 57*(10) pp. 1509–1528.

Sclater, N., & Bailey, P. (2015). Code of practice for learning analytics *(Online only). Retrieved, February 29, 2016, from: https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics*

Tsai, Y. S., & Gasevic, D. (2017). Learning analytics in higher education---challenges and policies: a review of eight learning analytics policies. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference (pp. 233-242)*. ACM.

University of Edinburgh (2017). Development of the University's Policy on Learning Analytics – Draft Policy Principles and Objectives (version 24 April 2017) *(Online only). Retrieved May 18, 2017, from: http://www.ed.ac.uk/files/atoms/files/learning_analytics_-_draft_principles_and_purposes_v24042017.pdf*

University of West London (2017). Learning Analytics Policy *(Online only). Retrieved May 18, 2017, from: http://www.uwl.ac.uk/sites/default/files/Departments/About-us/Web/PDF/uwl_learning_analytics_policy_redraft-1.pdf*

Measuring Student Self-regulated Learning in an Online Class

Qiujie Li
School of Education,
University of California, Irvine
qiujiel@uci.edu

Rachel Baker
School of Education,
University of California, Irvine
rachelbb@uci.edu

Mark Warschauer
School of Education,
University of California, Irvine
markw@uci.edu

ABSTRACT: Clickstream data has been used to measure students' self-regulated learning (SRL) in online courses, which allows for more timely and fine-grained measures as compared to traditional self-report methods. However, key questions remain: to what extent these clickstream measures provide valid inference about the constructs of SRL and complement self-report measures in predicting course performance. Based on the theory of SRL and a well-established self-report instrument of SRL, this study measured three types of SRL behaviors—time management, effort regulation, and cognitive strategy use—using both self-report surveys and clickstream data in an online course. We found both similarities and discrepancies between self-report and clickstream measures. In addition, clickstream measures superseded self-report measures in predicting course performance.

Keywords: Self-regulated learning, clickstream data, self-report survey, online learning.

1 INTRODUCTION

Perhaps even more so than in traditional in-person classes, students in online courses have great freedom to regulate their own learning by setting up goals and monitoring and controlling their cognition, motivation, and behavior to achieve their goals (Roll & Winne, 2015). Defined as self-regulated learning (SRL), these activities together play a critical role in determining students' success in online courses (Yukselturk & Bulut, 2007). Pintrich (1991) identified three main components of SRL: resource and effort management, cognitive strategy use, and metacognitive strategy use. Based on the framework, he developed a questionnaire, the Motivated Strategies for Learning Questionnaire (MSLQ), to measure SRL in in-person classes. Although the validity of MSLQ has been demonstrated in online classes (e.g., Yukselturk & Bulut, 2007), traditional self-report surveys taken over long intervals cannot capture how SRL unfolds nor can they provide timely measures to examine how SRL changes with environmental factors.

Researchers have been using clickstream data from online learning platforms to describe, interpret, and evaluate student SRL behaviors (Roll & Winne, 2015). However, very few of these studies are guided by educational theories in terms of how to develop and interpret clickstream measures (Gašević, Dawson, Rogers, & Gašević, 2016). Following MSLQ, we used both self-report surveys and clickstream data to examine three categories of SRL behavior in an online class: time management, effort regulation, and cognitive strategy use. Specifically, we investigated 1) to what extent clickstream measures provided valid inference about the constructs of SRL, 2) to what extent self-report measures reflected students' actual behaviors, and 3) to what extent clickstream measures complemented self-report measures in predicting student course performance.

2 METHODOLOGY

2.1 Data

Data were collected from a 10-week fully-online course offered by a public university in Fall 2016. The course contained four modules and each module was comprised of 9-14 small segments. The instructor recommended students complete one module every two weeks. A total of 319 students enrolled in this course. Self-report data was collected through pre- and post-course surveys. Clickstream data was collected through the course platform. The institution also provided data on students' demographics, prior academic achievement, and performance in the course.

2.2 Measures

2.2.1 Self-report measures

Measures updated from MSLQ were used both in the pre- and post-course surveys. Time management behaviors were measured by two statements: "I keep/kept a record of what my assignments are/were and when they are/were due" and "I plan/planned my work in advance so that I could turn in my assignments on time." Effort regulation was measured by seven statements (e.g., "I am/was quick to catch up with coursework when I started falling behind"). The use of cognitive strategies was measured by whether a student used elaboration and organization strategies regularly in the course.

2.2.2 Clickstream measures

Three clickstream measures were defined to measure time management. First, the clickstream measure of *studying a segment on time* was operationalized as a student visited segment page *i* before the deadline of segment *i*. For each module, the proportion of segments for which a student studied on time was calculated. Second, students might *postpone studying a segment until close to the deadline.* This behavior can be captured by the time interval between the deadline of segment *i* and the timestamp when a student visited segment page *i* for the first time. The average value of such time intervals for each module was calculated. Third, to meet the deadlines, students might *cram by studying a lot of segments over a short period of time*. This behavior can be captured by the standard deviation of the time intervals between the deadline of a segment *i* and the timestamp when a student conducted an initial visit of segment page *i*. The behavioral indicator of *quickly catching up after falling behind* was defined to measure effort regulation. It was operationalized as the time interval between the timestamp when a student conducted an initial visit of segment page *i* and the deadline of segment *i*. A clickstream measure of *reviewing a previous segment after studying a new one* was defined to capture students' use of the cognitive strategy of elaboration, since it indicated students attempting to link new ideas to knowledge already known. It was operationalized as a student re-visiting segment page *j* within T hours after the initial visit of segment page *i*.

3 PRELIMINARY RESULTS

The results section focuses on time management measures for which we now have preliminary findings. As the course progressed, students were less likely to study the segments on time and were more likely to postpone their study and study a lot of segments over a short period of time (see Table 1). In module 1, students visited, on average, 84% of the segment pages before the deadline. These numbers decreased for modules 2 and 3. On average, students visited the segment pages about 7 days before the deadline in module 1, but only about 4 days before the deadlines in the other three modules. Students were more likely to cram in modules 2, 3, and 4 than in module 1.

Table 1: Descriptive statistics of clickstream measures of time management

	Module 1		Module 2		Module 3		Module 4	
	M	SD	M	SD	M	SD	M	SD
Study on time	0.84	0.22	0.83	0.24	0.76	0.27	0.79	0.29
Postpone studying	7.19	3.12	3.97	2.84	3.79	3.27	4.40	3.13
Cram	3.66	1.43	2.43	1.51	3.06	1.89	2.78	1.57

3.1 Relationship between Self-report Measures and Clickstream Measures

The Pearson's correlations between clickstream measures and post-course survey measures were moderate, positive, and significant. The results provide evidence for the validity of the clickstream measures of time management behaviors. On the contrary, the correlations between clickstream measures and pre-course survey measures, in general, were small and insignificant, suggesting that students' anticipated behaviors may not reflect their actual subsequent behaviors in the course.

Table 2: Correlations between clickstream measures and self-report measures

Clickstream measures	Pre-course survey		Post-course survey	
	Keep record of deadlines	Plan in advance	Keep record of deadlines	Plan in advance
Study on time	0.09	0.08	0.21***	0.28***
Postpone studying	0.05	0.18**	0.21**	0.34***
Cram	0.06	0.1	0.21**	0.13*

Note. Average values of clickstream measures across four modules were used. * p<0.05 ** p<0.01 *** p<0.001

3.2 Relationship between Clickstream Measures and Achievement

We then regressed course grade on self-report and clickstream measures of time management, controlling for student background characteristics such as gender and high school GPA. In general, clickstream measures superseded self-report measures in predicting achievement. Specifically, both the clickstream measures of studying on time and postponing study predicted course grade. Self-reported measures on the pre-course survey did not predict course grade. Although the self-report measure of planning in advance on the post-course survey predicted course grade, the coefficient decreased and became insignificant after adding the clickstream measures.

Table 3: Time management measures predicting course grade

	Pre-Course Survey				Post-Course Survey			
	B	SE	B	SE	B	SE	B	SE
Keep record of deadlines	-0.017	0.17	-0.036	0.15	0.082	0.15	0.049	0.14
Plan in advance	0.088	0.19	0.035	0.17	0.16*	0.14	0.029	0.14
Study on time			0.28***	0.86			0.26**	0.89
Postpone studying			0.20**	0.06			0.18**	0.06
Cram			0.006	0.15			0.007	0.15
R^2	0.21		0.40		0.25		0.40	

Note. All coefficients are standardized. * p<0.05 ** p<0.01 *** p<0.001

4 FUTURE WORK

The preliminary results suggest that clickstream measures of SRL may offer more insightful and valid information about students' actual learning processes than their own anticipations or perceptions. In the next phase of research, we will employ other algorithms, such as classification trees, to identify nonlinear relations between measures of SRL and achievement, and search for interactions between different measures in predicting achievement.

REFERENCES

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68-84. https://doi.org/10.1016/j.iheduc.2015.10.002

Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). Retrieved from http://files.eric.ed.gov/fulltext/ED338122.pdf

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, *2*(1), 7-12. http://dx.doi.org/10.18608/jla.2015.21.2

Yukselturk, E., & Bulut, S. (2007). Predictors for student success in an online course. *Journal of Educational Technology & Society*, *10*(2).

Categorization, Intersectionality, and Learning Analytics

Timothy A. McKay
University of Michigan
tamckay@umich.edu

Galina Grom
University of Michigan
grom@umich.edu

Benjamin Koester
University of Michigan
koester@umich.edu

ABSTRACT: Learning analytics often relies on data produced by education systems which include traditional categorical descriptors of identity. Uncritical use of these reductive categories obscures the complexity of identity and masks the unique experience of each student. If learning analytics is to accomplish its goal of understanding and improving teaching and learning for all students, it must examine the methods it uses to account for social identity more closely. In this work, we describe how feminist studies of intersectionality have informed our own analysis of how social identity might influence student performance in an array of large introductory courses.

Keywords: Social Identity; Categorization; Intersectionality; Personalization

1 INTRODUCTION

Data sets used in learning analytics regularly record categorical descriptors of each individual's identity: gender, underrepresented and first generation status, residency, race. Analyses based on these descriptors often proceed along individual dimensions; comparing male and female, first and non-first-gen, or racial and ethnic categories. Such analyses elide over the lived experience of identity, which is neither simply categorical nor unidimensional. This reality, long recognized by those who study social identity, is often described as intersectionality (Davis, 1981; Crenshaw, 1989). This use of reductive categorization to describe complex individuals is a persistent problem in the world of big data. Cheney-Lippold (2017) refers to these categories as "measureable types", distinguishing (for example) between gender as a lived experience and 'gender' as a label within a data set. In this work, we will adopt his convention, denoting the simple, transcoded measurable types which stand in for complex social identities by enclosing them in single quotes.

If we are to fully realize the ambition of learning analytics, "to understand and optimize learning and the environments in which it occurs" (Siemans & Long, 2011), we must move beyond the information loss associated with the use of measureable types and strive to characterize the individuals we study in a holistic, multidimensional way. In this brief research abstract, we describe essential elements of our efforts to move beyond the reductive characterization of our learners. We begin with an overview of methodological approaches to dealing with intersectionality. This is followed by a concrete example, based on efforts to understand gendered performance differences in large introductory science courses. We conclude with some lessons learned and a vision for using analytics for deeper personalization at scale.

2 METHODOLOGIES FOR ADDRESSING INTERSECTIONALITY

There are many approaches to confronting intersectionality when attempting to understand the relation between social identity and subject formations. McCall (2005) provides a useful framework for considering the range of possibilities, though we recognize the irony of discretely classifying methods for addressing intersectionality.

Anticategorical Complexity asserts that social life are irreducibly complex, and that categories imposed on them usually exist to produce and enforce inequalities. Nonreductive, this approach is best able to capture the full complexity of each individual's social identity. It is, in a sense, a demand for absolute personalization: for seeing each individual only as an individual, not a member of any collective category. *Intracategorical Complexity* aims to explore the deeper, often hidden diversity which exists within a cell (or cells) of traditionally constructed categories. It focuses on deconstructing the apparent homogeneity of a data category, critiquing the uncritical use of measureable types so common in data science. Finally, *Intercategorical Complexity* takes traditional categories as provisional, and uses them to frame analyses aimed at documenting existing "relationships of inequality" among these groups.

3 INTERSECTIONAL EXPLORATIONS OF ACADEMIC PERFORMANCE

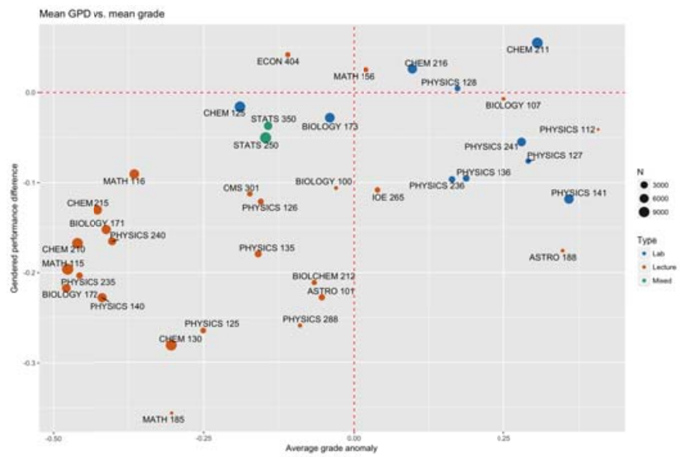
To illustrate the importance of an intersectional approach in learning analytics, we provide a concrete example – studies of gendered performance difference in large introductory science courses. In this work, we compare the performance of each student to a simple expectation – their performance in all other courses completed at the same institution. We call this difference between GPA in all other classes and course grade a student's 'grade anomaly'; a course-and-student-specific measure of better or worse-than-expected performance. Courses which award comparatively low grades have average grade anomalies (AGAs) which are negative; those which award comparatively high grades have AGAs which are positive. In this approach, the performance of two groups of students may be compared by examining the difference in AGA for the two groups.

Our studies began as an effort to explore gendered performance differences (GPD = AGA_{female}-AGA_{male}) in individual courses. They have since expanded to examine overall AGA and GPD across a wide range of large introductory courses. Figure 1 plots GPD vs. AGA for an array of 37 large introductory courses in science, engineering, and economics. This figure shows that while lecture courses exhibit large gendered performance differences, lab courses typically do not. These GPDs are persistent over many years and independent of instructor identity. They cannot be explained by reference to any other prior information available in our student record system (Blinded internal study, 2016).

We interpret these unexplained GPDs as signs of structural inequity in these courses and are currently exploring several approaches to eliminating them. Since the courses are unusual in their reliance on high stakes, timed examinations for determining grades, we suspect that stereotype threat associated with social identity may play a role in the creation of these inequities. This possibility makes our understanding of social identity especially important in this context, and has driven us to investigate our use of traditionally constructed categories in characterizing students.

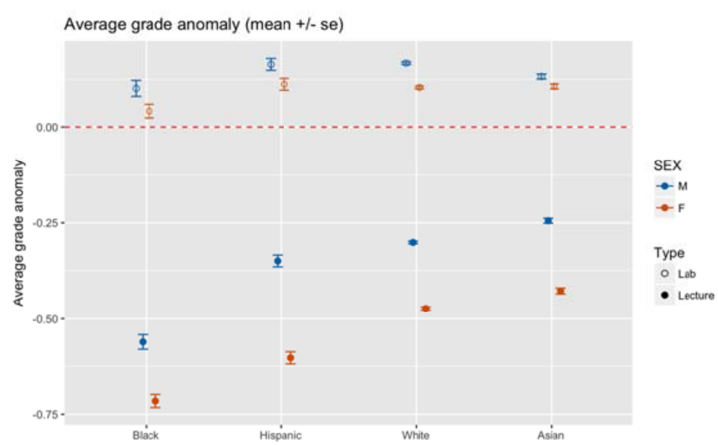
Our original analysis relies on a classic binary characterization of student identity through a single measureable type – 'male' or 'female' – contained within our student record system. The widespread and persistent appearance of GPDs demonstrates that the label 'female' is correlated with underperformance in these courses, but it does little to reveal what about students actually *causes* underperformance.

Figure 1: Gendered Performance Diff. is plotted vs. Average Grade Anomaly for a series of 37 lecture, lab, and mixed format courses across a range disciplines. While lecture courses show substantial GPDs, lab courses do not.



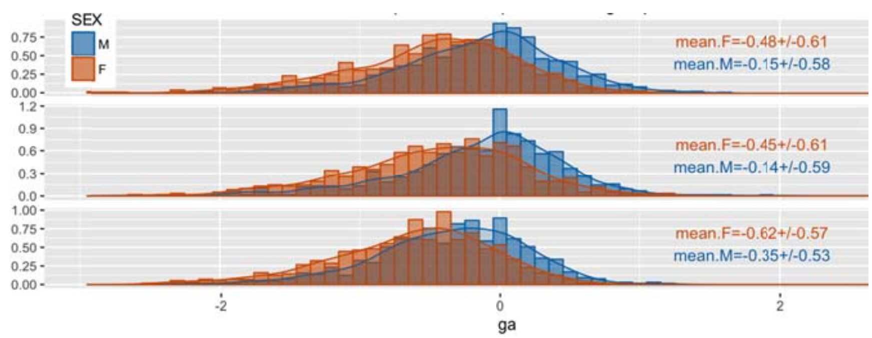
An application of the *Intercategorical Complexity* approach to better and worse-than-expected performance might begin (for example) by examining the intersection of 'gender' and 'ethnicity'. Figure 2 shows the results of such an examination, comparing AGA in an array of biology, chemistry, physics, and economics courses for male and female students in four 'ethnicity' groups. Clearly 'gender' is not the only factor significally affecting AGAs. Indeed 'Black male' students underperform relative to 'White female' and 'Asian female' students. In many analyses, these potentially intersecting identities are treated as independent. In fact they may interact, creating non-linear effects. If we pursue solutions to student performance gaps with *only* the lens of 'gender', we would miss essential elements of the student experience.

Figure 2: Average grade anomalies are shown for both 'male' and 'female' students in four 'ethnicity' groups in six years of data drawn from a series of biology, chemistry, physics and economics lecture and lab courses.



To apply the *Intracategorical Complexity* approach to this analysis, we take one category of students – 'male' or 'female' for example – and use additional information to probe the correlates of underperformance within it more deeply. To illuminate this possibility, Figure 3 shows the distribution of grade anomalies for individual ‘male’ and ‘female’ students in one of these courses over three successive fall terms. While a statistically and materially significant difference in average grade anomaly clearly exists, there is enormous overlap among the outcomes of individuals, with many ‘male’ students performing worse-than-expected and many ‘female’ students performing better-than-expected. This substantial variation within measureable types drives us to consider more closely what other, perhaps unmeasured factors might be responsible for underperformance. Are there some ‘female’ students immune from social identity threat; some ‘male’ students subject to it?

Figure 3: Shows grade anomaly distributions for all ‘male’ and ‘female’ students in 3 years of data from a large lecture course.



The ultimate goal might be to approach the data with the lens of *Anticategorical Complexity*, resisting the temptation to reduce the unique social identities of students to traditional categories of measureable types entirely. When we speak of true personalization at scale, this, we believe, should be the goal: to spend most analytic effort understanding and improving the experience of individuals, rather than ending our analysis with an array of traditionally constructed measureable types.

4 IMPLICATIONS

It is common to characterize students using only the traditionally constructed measureable types: researchers are inclined to use the data they have. If the learning analytics community is going to achieve its goals, it will have to maintain a rigorously critical stance toward traditional measureable types. The feminist literature of intersectionality provides important insight into how this work might proceed, and those using data to understand and improve teaching and learning have much to learn from it.

REFERENCES

Koester, B., Grom, G. and McKay, T. (2016). *Patterns of gendered performance difference in introductory STEM courses.* arXiv preprint arXiv:1608.07565.

Cheney-Lippold, J. (2017). *We Are Data: Algorithms and The Making of Our Digital Selves*. NYU Press.

McCall, L. (2005). The complexity of intersectionality. *Signs: Journal of women in culture and society*, *30*(3), 1771-1800.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, *46*(5), 30.

Practitioner Papers

Using Knowledge Tracing to Tune an Existing Student Teaching Tool

Author(s): Neil L Zimmerman
McGraw-Hill Education
neil.zimmerman@mheducation.com

ABSTRACT: This paper documents a research project designed to validate an existing learning tool, including the methodology, the findings, and the changes we are making to this product as a consequence of these findings. SmartBook is an educational suite that is used by college students to quickly master topics that appear in a textbook. This software uses a complex and proprietary algorithm to decide whether students have achieved adequate mastery of a topic. While student satisfaction with this software is high, we wanted to quantify the extent to which students learn using this software to ensure that students are learning as much as is hoped. To do this, we implemented Performance Factors Analysis and Bayesian Knowledge Tracing at scale using Apache Spark, measured their ability to describe learning and predict student performance in SmartBook. We then used our preferred model to estimate when and if the student mastered the material. Investigation of the modeled outcomes reveals a straightforward modification that we are making to the next version of SmartBook, which is briefly discussed.

Keywords: **• Information systems → Data mining; • Applied computing → E-learning; • Computing methodologies → Supervised learning by regression; • Mathematics of computing → Markov networks**

1 INTRODUCTION

As part of our work on SmartBook, a web application used primarily in a higher-education setting, the author of this paper needed to determine how well this application was working and make changes to the next version. In normal usage, instructors assign a set of Knowledge Components (KCs) to the class,[1] which correspond to a given passage or set of passages in the course reader, and also to a set of questions which will be administered by SmartBook. Students are expected to work through the assignment in SmartBook after they have done their course reading; the software will determine whether students have mastered each KC, encourage them to review the sections corresponding to KCs for which they have not demonstrated sufficient mastery, and give them more questions until they have been deemed to have sufficiently mastered all the assigned KCs.[2] Currently, SmartBook uses a set of heuristics to determine mastery that do not lend themselves to quantitative analysis, so in order to measure how well our platform was teaching, we first had to choose and implement a framework to decide how much students had learned.

[1] In designing both the content and the teaching software, we treat KCs as equivalent to the skills or concepts one would find in a q-matrix. However, systematically testing this assumption remains the subject of future research.

[2] In this paper, we call the appearance of a question associated with a given KC an *instance* of the KC.

SmartBook is not an Intelligent Tutoring System (ITS): the interventions available to students who have not acquired a Knowledge Component are effectively limited to referring them to the relevant passage of the course reader. No assumptions are made about the cause of incorrect answers, and no KC prerequisite graph exists. On the other hand, KCs are granular and independent, analogous to the production rules that would be found in an ITS (Corbett et al. 1997). Consequently, we believe that it is appropriate to apply traditional knowledge tracing techniques, which use machine learning to infer a student's performance on upcoming instances of a KC from their performance on previous instances, to determine and describe student mastery of each KC.

This paper documents our analysis of a one-semester course in SmartBook, which was administered in many schools hundreds of times over the period of about five years; the logs for this course encompass roughly 90,000 students, 6,000 KCs, and 40,000,000 distinct interactions with the software.

2 METHODOLOGY

SmartBook is broadly intended to teach students the things the instructor decides they need to know as quickly and with as little frustration as possible; this does not lend itself to a single quantitative evaluation metric. For this revision of the software, we decided to focus on whether students end with sufficient mastery of each KC they are assigned; this is both the most important goal of the software and the easiest to infer from student interaction logs.

The most natural approach, in our view, was to use knowledge tracing to describe mastery in terms of the probability that a student would respond correctly to an instance of a KC if they saw that KC again. Using the open source Apache Spark, we implemented parallel versions of two classic knowledge tracing models, Bayesian Knowledge Tracing (BKT) (Corbett and Anderson 1995) and Performance Factors Analysis (PFA) (Pavlik et al 2009). While our implementation followed those papers as closely as possible, implementing these in Spark involved rewriting most of the code from scratch. Because PFA is essentially a logistic regression with some feature engineering, we implemented it entirely in Spark MLLib. Spark contains no Hidden Markov Model implementation, so for our BKT routine we used Spark as a parallelization framework and used the Pomegranate package[3] to implement one BKT model for each KC.

We randomly divided 70% of our students into our training set and 30% into our testing set, and evaluated both models on their ability to predict correct or incorrect responses in the testing set. We found that both models performed very similarly: BKT had an area under the Receiver Operating Characteristic curve (AUCROC) of 0.761, and PFA had an AUCROC of 0.741. These numbers are consistent with previous research that used the same models to predict performance on ASSISTMENTS (Gong et al. 2011, Wilson et al. 2016), so we feel justified in concluding that student learning in SmartBook can be adequately described by existing models.

While PFA and BKT perform similarly well at predicting correct answers, they do not describe student learning in the same way. PFA makes no attempt to model the probability that students

[3] https://github.com/jmschrei/pomegranate

guess correctly, or slip when they possess the required knowledge: the only probabilities it produces are the probability of a correct answer, which the researcher must decide how to interpret in terms of knowledge. BKT, on the other hand, models a hidden state, whether the student knows the KC or not, and also the probability of guessing and slipping to get a probability of a correct answer. Since we have every reason to believe that BKT is a valid model for describing learning in SmartBook, we decided to use its hidden state to describe the probability that a learner knows the KC.

3 RESULTS

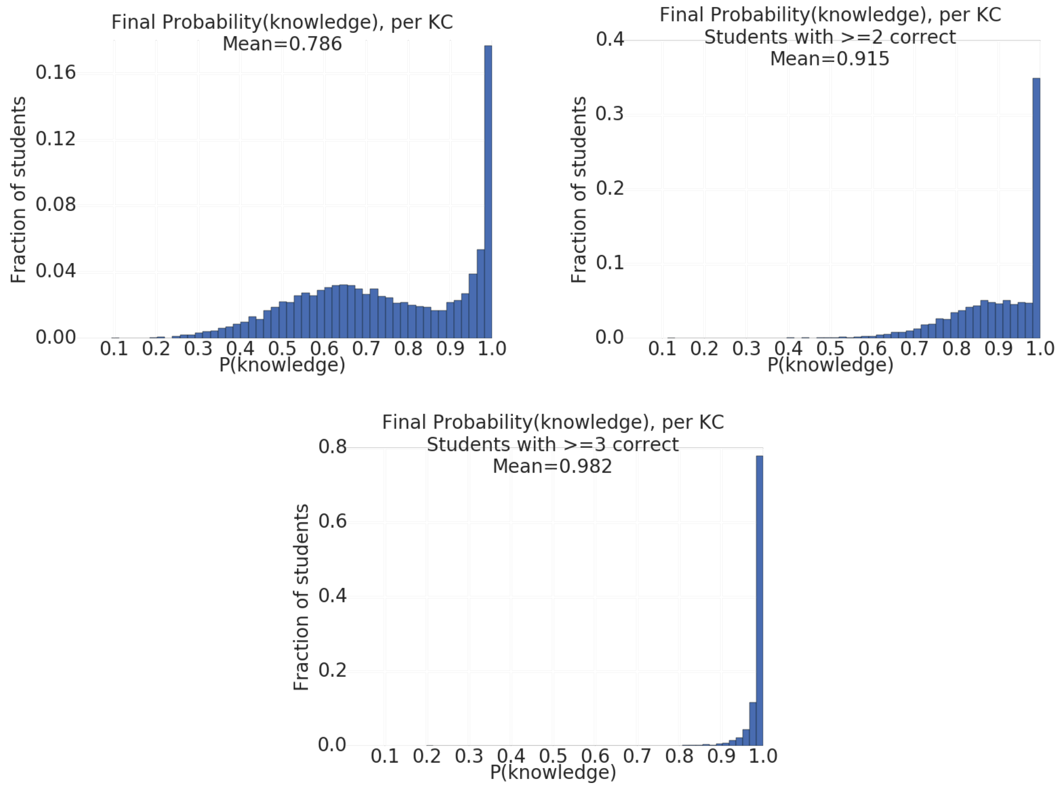


FIGURE 1: Top left, the probability that a student knows (per hidden state of BKT model) a given KC after the last time they see it, averaged across all students per KC. Top right, considering only students who have given two or more correct responses to a given KC. Bottom, considering only students who have given three or more correct responses. Means are across all student/KC interactions.

Because SmartBook is intended to ensure that students have mastered the material regardless of how much time they need, we concern ourselves primarily with the knowledge students possess when they finish being taught. Consequently, we evaluate the effectiveness of SmartBook based on the probability that the student knows a given KC after the last time they see it, and not the rate at which students learn.

The top left panel of Figure 1 shows these probabilities, per KC, averaged across all students who were assigned and attempted a given KC. Across all KCs and students, there is a 79% mean probability in the BKT model that a student will know a KC when they are done being taught it, versus a 16% probability that the students know a KC before they are taught it (not shown). Since

SmartBook is used in support of classroom learning, and students are normally assigned KCs in SmartBook before they are taught in class, we feel that a 79% probability of final knowledge does not necessarily represent a bad outcome for students, but we would like to do better if we can.

We believe that the lower-than-desired mean probability of knowledge arises from the fact that the SmartBook completion algorithm, as currently written, will under some circumstances consider a student to have sufficiently mastered a KC after only a single correct answer, if that correct answer comes after the student has gotten a certain number of incorrect answers and has gone back to review the material. In the top right panel of Figure 1, we see that after two correct answers, students are much likelier to know the KC (92%), and after three correct answers they are nearly certain to know it (98%).

4 CONCLUSIONS AND FUTURE RESEARCH

In light of this finding, the authors of this paper decided to make a slight modification to the KC completion algorithm in SmartBook to require that students always provide two correct answers on a KC before they are marked as having completed it. The efficacy of this modification will be validated over the course of the upcoming years.

SmartBook questions are continually being refreshed; before the next revision we intend to use a model like KT-IDEM (Pardos & Heffernan 2011) to check whether instances of KCs have similar guess and slip terms; in principle this should be the case, but using a model to prove or disprove this will allow us to make better authorial decisions.

REFERENCES

Corbett, Albert T., and John R. Anderson. "Knowledge tracing: Modeling the acquisition of procedural knowledge." *User modeling and user-adapted interaction* 4.4 (1995): 253-278.

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. *Handbook of human-computer interaction*, *5*, 849-874. https://doi.org/10.1016/b978-044481862-1/50103-5

Gong, Y., Beck, J. E., & Heffernan, N. T. (2011). How to construct more accurate student models: Comparing and optimizing knowledge tracing and performance factor analysis. *International Journal of Artificial Intelligence in Education*, *21*(1-2), 27-46.

Pardos, Z., & Heffernan, N. (2011). KT-IDEM: introducing item difficulty to the knowledge tracing model. *User Modeling, Adaption and Personalization*, 243-254.

Pavlik Jr, P. I., Cen, H., & Koedinger, K. R. (2009). Performance Factors Analysis--A New Alternative to Knowledge Tracing. *Proceedings of the 14th International Conference on Artificial Intelligence in Education*

Wilson, K. H., Xiong, X., Khajah, M., Lindsey, R. V., Zhao, S., Karklin, Y., ... & Heffernan, N. (2016). Estimating student proficiency: Deep learning is not the panacea. In *Neural Information Processing Systems, Workshop on Machine Learning for Education*.

Do Students Notice Notifications? Large-Scale Analysis of Student Interactions with Automated Messages about Grade Performance

Presentation Submission

Author(s): John Whitmer, Ed.D.
Blackboard Inc.
john.whitmer@blackboard.com

Daniel Nasiatka
Blackboard Inc.
daniel.nasiatka@blackboard.com

Timothy Harfield, Ph.D.
Blackboard Inc.
timothy.harfield@blackboard.com

ABSTRACT: Despite widespread interest in educational technology features providing data to students, there are few research studies providing evidence of how students respond to these features, let alone their efficacy in improving learning outcomes. In this study, we analyzed clickstream data from student activity in relation to automated notifications provided to students in a Learning Management System (LMS), specifically messages about grade increases and decreases over time, as well as student grades compared to other students in the course. A data set extracted from the LMS with 18,727 notifications sent to 2,592 students in 405 courses at four institutions was used. Students opened notifications at relatively high rates overall, and demonstrated a clear preference for notifications that compared them to peers in their course over notifications about trends in their own activity. Students were clustered into five groups based on the types of notification received. The clusters indicated that student behavior within the LMS is highly consistent over the duration of a course. Open rates were different between the student clusters, yet had some consistent patterns. These rates and patterns suggest different underlying student motivations and study habits.

Keywords: Student analytics, data dashboards, self-regulated learning, educational technology, learning management systems

1 INTRODUCTION

Providing students with information about how they are performing in a course, and alerting them in advance if they are at risk of not passing a course, has been suggested as a powerful way that learning analytics can be used to improve educational outcomes (Dahlstrom, Brooks & Bichsel, 2014). Building on prior research demonstrating significant relationships between LMS use and student grades (Rafaeli and Ravid 1997, Macfadyen and Dawson 2010, Fritz 2011), educational technology vendors are building student-facing dashboard functionality within applications, providing direct interventions to the people who need them most and in ways that are immediately actionable. However, there is little empirical research in this area. Research that has been conducted has involved relatively small samples of students and has reported mixed results – including some suggestions that this type of information could demotivate students at risk of failing a course (Aguilar, 2016).

One concern raised in prior research was whether all students would benefit from the same message and type of analytic feature, or whether these features should be differentiated based on student background characteristics and performance within a course (Teasley, 2017). However, integration with student information systems and other data sources is technically difficult, and educational technology vendors would create barriers to adoption if these data sources were required before any notifications could be created. There is a compelling practical need to provide notifications with the information solely available within the Learning Management System (or other educational technology platform), which does not provide information about student educational experience, demographic characteristics, or other contextual information.

Therefore, in a new LMS user experience, a feature was created to send students (and faculty) notifications about individual trends in performance and performance relative to peers, using rule-based thresholds to identify high and low performers. A prior research study found that students from low GPA backgrounds valued this information more than students from high GPA backgrounds (Teasley & Whitmer, 2017). This prior research was conducted through interviews and surveys of students, asking them to reflect on their experience in prior classes. While these results were instructive and supported our design approach, we wanted to cross-reference them to actual student behavior. The feature has been deployed, and we now have data from real instructional contexts that can be analyzed.

The following questions oriented this study:

1. Do students consistently receive the same notifications? Can the notifications be used to classify student performance?

2. Are students interested in notifications, as indicated by the rate at which they open them?

3. Does this interest vary by the type of notification received or student classification?

2 CONTEXT

Rule-based notifications were added to the new user experience of an LMS to encourage students to improve their performance and recognize positive performance and LMS activity. The notifications use logic rules that compare students to their behavior in the prior week (e.g. trends in their behavior) as well as to that of other students in the course. There were four notifications investigated in this study: high and low performance relative to peers (top 10% and bottom 5% course grade, "GradeinHighest" and "GradeinLowest"), as well as increases and decreases in behavior relative to the same student's performance in a prior week (increase or decrease of 10%, "GradeIncreased" and GradeDecreased"). The thresholds used were based on empirical research of student grade distributions within the LMS. Rule names and thresholds were used internally; the student experience was transformed through design treatments.

The messages used in the notifications are intentionally "light," using a "concerned friend" tone to keep students' interest while providing sometimes concerning information about the behavior or performance in a course. A screenshot of a positive message illustrating the first notification in this workflow is provided in Figure 1.

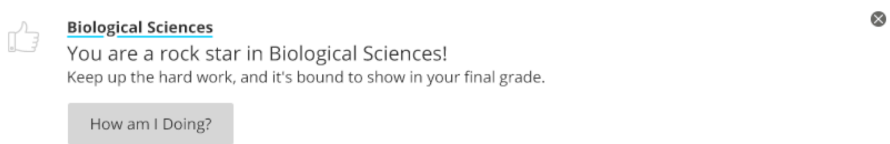


Figure 1: Positive Notification Message Example

Notifications were rendered within the LMS "activity stream," a centralized interface that provides students and faculty with information about their courses: assignments due, other deadlines, etc. If students click a notification, they are presented with a dashboard that shows more detailed information. This is usually a chart that plots their performance over time or illustrates their position relative to other students in the course. Finally, there is a follow-on action that a student is encouraged to take to address the notification. This pattern of interactions and workflow is consistent for all notifications in the course.

3 DATA SAMPLE AND METHODS

Item	N
Notifications	18,727
Students	2,592
Courses	405
Institutions	4

Table 1: Data Set

In this study, we used a sample of courses running in this LMS (hybrid and fully online) with the new design feature enabled. We performed an archival investigation (no A/B testing) with data collected on notifications throughout a semester. We filtered the institutions and data sample to include institutions with a large number of notifications (n>2,000) and courses with more than three average notifications per student. Filtering still included a large sample, as described in Table 1, from which findings can be generalized to a larger population. The data set used was anonymized. No student, faculty, nor course-level descriptive information was included in the data set.

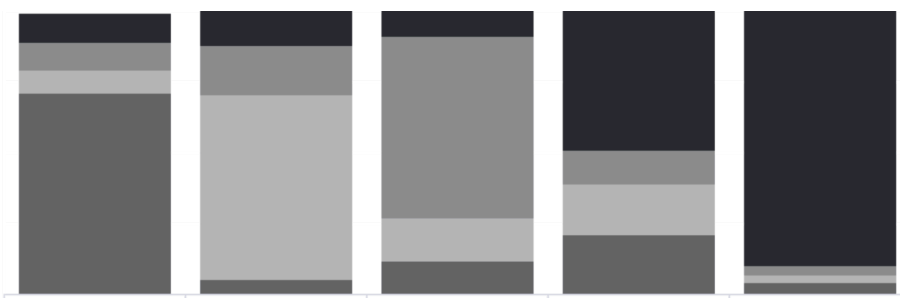
Several statistical methods were applied to the data set to answer the research questions. First, students were grouped by patterns in the types of notification received. For this analysis, we implemented k-means cluster analysis. The number of notifications per type per student was aggregated, as the goal was to cluster students across all courses. The input to the algorithm was the normalized number of notifications received in each category (as determined by the percentage of each notification that a student received). To perform the statistical significance tests for the difference in notification open rates between groups, we implemented chi-square tests with

Bonferroni adjusted p-values for multiple comparisons. The data was first aggregated to a granularity of notification type for an initial analysis; for a more detailed analysis, we aggregated the data to a granularity of student cluster and notification type. We considered all notifications rendered on student devices (i.e. computer, table, smartphone). If a student opened a notification multiple times, we only considered the first opening in our analyses of whether a notification was opened.

4 FINDINGS

The first item we investigated was the frequency with which students received the different types of notifications. This was important as a first step in order to pursue the question of whether students with different academic performance levels have different responses to the notifications. Prior research predicted that students with low performance were likely to open the notifications more frequently (Teasley & Whitmer, 2017), however, we wondered if students who consistently received high grade notifications were more active overall and would be more likely to open the notifications.

We identified five distinct groups. There is one cluster per notification type, with an additional type (the Malleable Middle) that receives a variety of notifications.



	Low Performers (391 students)	Decliners (303 students)	Improvers (230 students)	Malleable Middle (712 students)	High Performers (789 students)
GradeInHighest	10%	13%	9%	49%	91%
GradeIncreased	10%	17%	64%	12%	3%
GradeDropped	8%	66%	15%	18%	3%
GradeInLowest	71%	5%	12%	21%	4%
Total Notifications	*3,589*	*1,487*	*1,195*	*5,844*	*8,451*

Figure 2: Notification Type Distribution by Student Cluster

The count and distribution of cluster types are shown in Figure 2. The clusters have the following characteristics and implications about the students receiving these notifications:

1. High Performers almost exclusively received notifications that they had a high grade relative to other students in the course. These students started with high grades and continue to earn high grades. The largest proportion of students were in this category, perhaps due to the 10% rule threshold.

2. Low Performers were in the opposite position from High Performers. They almost exclusively received notifications that they had a low grade relative to other students. They received

more of other notification types than High Performers, but this was by far the most frequent notification that they received.

3. Improvers largely received notifications that their grade had increased. These are students whose grade increased throughout the term, but rarely to the extent necessary to place them in the top 10% of the class.

4. Decliners were in the opposite position as Improvers. Their grade started high but decreased frequently in the course. These students also received other notification types, but rarely received GradeInLowest. It is possible that these students started with a very high grade and declined yet still passed the course.

5. Students in the Malleable Middle show similar traits to High Performers, with over twice as many notifications in the GradeInHighest type compared to other types, but they also received a large number of GradesDropping and GradesInLowest notifications. These students had highly varied achievement during a course.

Given the different number of notifications between clusters, there may be concern about the validity of using percentage of notifications as data for clustering. However, this analysis was not sensitive to filtering for students receiving a high or low volume of notifications. These clusters were also consistent across institutions. In the subsequent analyses, we examine how each of these groups interacts with the LMS notifications that they received.

The clusters of students indicate that there were very different profiles of students in the study in terms of academic performance, and likely their interactions with course materials and activities, such as notifications, would also be different. The notification open rates by cluster are illustrated in Table 2. As is evident in this figure, there are substantial differences in overall open rates by cluster as well as the open rates by notification type within each cluster.

Table 2: Open Rates (all notifications) between Student Clusters

Cluster 1	Cluster 2	N1	N2	Open Rate 1	Open Rate 2	P-Value
Decliners	High Performers	1487	8451	28%	44%	2.62E-31
Decliners	Improvers	1487	1195	28%	41%	1.33E-12
Decliners	Low Performers	1487	3589	28%	34%	6.16E-06
Decliners	Malleable Middle	1487	5844	28%	35%	5.10E-08
High Performers	Improvers	8451	1195	44%	41%	7.04E-02

High Performers	Low Performers	8451	3589	44%	34%	1.18E-21
High Performers	Malleable Middle	8451	5844	44%	35%	5.99E-24
Improvers	Low Performers	1195	3589	41%	34%	5.30E-05
Improvers	Malleable Middle	1195	5844	41%	35%	2.74E-04
Low Performers	Malleable Middle	3589	5844	34%	35%	3.73E-01

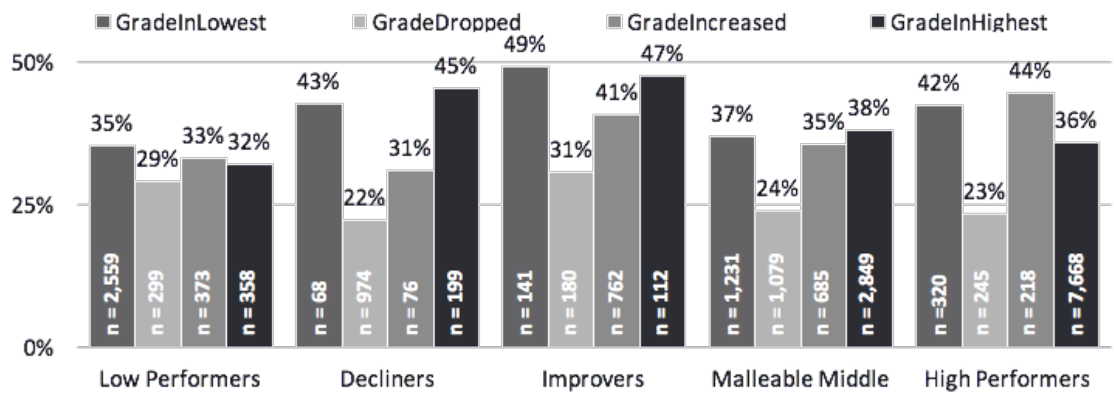


Figure 3 - Notification open rates by student cluster and notification type

The open rates were different between most pairs of student clusters, with substantial differences between types of notifications, as shown in Table 2 and Figure 3. High Achievers showed the highest open rates overall (44%), followed by the Improvers (41%). There was not a statistically significant difference in open rates between the High Performers / Improvers and between the Low Performers / Malleable Middle; these were the relative high and low open rate pairs. These findings illustrate that it is not sufficient to look at overall open rate differences, as the characteristics of students receiving these notifications vary widely. These findings are also consistent with the research mentioned previously that found a clear relationship between LMS use and student achievement.

Another finding was that the cluster of students receiving varied notifications (the Malleable Middle) tended to have lower interaction with the notifications. Although these students most often received GradeInHighest notifications (similar to High Performers), their open rates were not significantly different from the Low Performers. This is a surprising result. It would seem that novel and diverse notifications would be of more interest to students than a steady stream of the same message, but that was not the result in practice.

Significant differences were found within each cluster of students between the relative-ranking notifications (GradeInHighest and GradeInLowest) and the grade change notifications (Grade Increased and Grade Dropped) for all clusters except for Low Performers ($p < 0.01$). This finding mirrors what was discussed previously for the overall value of these types of notifications and those explanations apply equally to these analyses. However, it is notable that Low Performers did not show a significant difference in open rates between comparative information and self-information notifications. Furthermore, Low Performers show significantly less interest in all notifications than other groups. This may indicate that these students have lower overall activity and interaction with the LMS than other groups. An alternative explanation is that these students are more sensitive to constructive information, as they almost exclusively received messages about low performance.

It is also interesting to note that none of the student clusters showed a statistically significant difference in open rates for positive/negative feedback for relative-ranking notifications (GradeInHighest vs GradeInLowest), while several groups did show a significant difference in open rates for positive/negative feedback for grade change notifications (GradeIncreased vs GradeDropped). That is, the results indicate that the students interested in comparative information are indifferent to whether the feedback is positive or negative, while students show a significant preference for positive feedback when presented with a notification focusing on the student's own performance. As the results from the prior study with students showed (Teasley & Whitmer, 2017), this finding could suggest an opportunity to refine the messaging in notifications in order to engage students better. Alternatively, another explanation could be that students of all types of students are more sensitive to negative personal feedback than negative comparative feedback.

5 LIMITATIONS AND NEXT STEPS

This study investigates how students receive and interact with automated notifications based on their course grades. The ultimate question is whether this behavior has a noticeable impact on student grades or online activity. We attempted to investigate this question for this study but encountered confounding factors and unclear data that led us to believe that an archival study of this question at scale is not feasible. Students in this data set often received multiple types of notifications from courses around the same time, so it was difficult to attribute any outcome to a single notification. In addition, there were large variations in course grades and activity between groups and over time. A better approach to exploring this relationship would be to conduct a small-scale study with a single course or group of courses; ideally using an experimental design or longitudinal analysis.

This study was also limited to investigation of grade-based notifications and a relatively small number of institutions. Future research should expand into LMS behavior notifications and a larger number of institutions.

6 CONCLUSIONS

In this study, we contribute to the research literature and communities of practice interested in student-facing learning analytics notifications. The findings indicate that students have strong interest in this type of information to assist them with their academic endeavors. This research differs from prior work in that we use anonymized archival data collected from authentic courses that provides insights into actual student interactions and behavior with these notifications.

In most cases, students exhibited consistent trends in their grade achievements that resulted in them receiving the same notifications over time. The distribution is partially due to the thresholds used for the notifications in the study, but also indicates that student grade achievement and position relative to other students is largely consistent over time. Surprisingly, students still have a strong interest in the notifications that they receive, despite receiving a consistent message that is associated with their course position.

Students also demonstrated a clear preference for notifications comparing them to other students over notifications showing changes in their achievement over time. This is a key advantage of these types of notifications and learning analytics dashboards. Student responses confirm student interest in this type of feature. Further, students appear to be most interested in notifications that recognize positive achievement compared to those that identify areas for improvement, counter to the intent of most notifications, which is to identify at-risk students and help motivate changes in behavior. These approaches can co-exist, but this finding emphasizes the importance of keeping positive nudges in notifications that are developed.

REFERENCES

Aguilar, S. (2016). Perceived motivational affordances: Capturing and measuring students' sense-making around visualizations of their academic achievement information. (Doctoral Dissertation) University of Michigan, Ann Arbor, MI.

Dahlstrom, E., Brooks, D.C., and Bichsel, J. (2014, September). The Current Ecosystem of Learning Management Systems in Higher Education: Student, Faculty, and IT Perspectives. Research report. Louisville, CO: ECAR. Available from https://net.educause.edu/ir/library/pdf/ers1414.pdf.

Fritz, J. (2011). "Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers." The Internet and Higher Education 14(2): 89-97.

Macfadyen, L. P. and S. Dawson (2010). "Mining LMS data to develop an "early warning system" for educators: A Proof of Concept." Computers & Education(54): 11.

Rafaeli, S. and G. Ravid (1997). OnLine, Web Based Learning Environment for an Information Systems course: Access logs, Linearity and Performance. ISECON 1997, Orlando, FL.

Teasley, S. D. & Whitmer, J. (2017). Surprising Lessons from Research on Student Feedback about Data Dashboards. Blog Post. Published February 2, 2017. URL: http://blog.blackboard.com/research-student-feedback-data-dashboards/

Teasley, S. D. (2017). Student Facing Dashboards: One Size Fits All? Technology, Knowledge and Learning. doi:10.1007/s10758-017-9314-3.

Evaluating the Adoption of a Badge System based on Seven Principles of Effective Teaching

Chi-Un Lei, Xiangyu Hou, Donn Gonda
Technology-Enriched Learning Initiative, University of Hong Kong, Hong Kong
{culei, hxiangyu, dgonda}@hku.hk

Xiao Hu
Faculty of Education, University of Hong Kong, Hong Kong
xiaoxhu@hku.hk

ABSTRACT: Badge systems are useful teaching tools which can effectively capture and visualize students' learning progress. By gamifying the learning process, the badge system serves to improve students' intrinsic learning motivations, while adding a humanistic touch to teaching and learning. The implementation of the badge system and the evaluation of effectiveness should be guided by pedagogical principles. This paper evaluates the effectiveness of a badge system in a non-credit-bearing outreach course from a pedagogical point of view based on Chickering's "Seven Principles for Good Practice in Undergraduate Education" and Object-Action Interface model. Furthermore, usage of the badge system is analyzed in terms of system traffic and the distribution of earned badges. The future development plan of the badge system is outlined. It is hoped that the findings in this paper will inspire teachers and e-learning technologists to make effective use of badge systems and other learning visualization tools for teaching and learning.

Keywords: learning visualizations, badges, learning design, case studies, use and evaluation

1 INTRODUCTION

Badges and learning progress dashboards have been adopted for tracking students' learning progress (Verbert et al. 2014; Auvinen et al. 2015; Mah et al. 2016). Teachers can use them for predicting students' final scores or dropout rates, as well as identifying at-risk students. Meanwhile, students can use them to monitor their own learning. Besides generic learning dashboards, gamified student-centric badge systems have been deployed in various institutions and learning contexts (Ian O'Byrne et al. 2015; de Freitas et al. 2017). The badge system is an effective visualization tool which provides a clear description of students' learning progress. Furthermore, by gamifying the learning process, the badge system serves to improve students' intrinsic learning motivations, while adding a humanistic touch to teaching and learning via badges, learning progressions, and learning quests.

Various studies on the use of badge systems and dashboards in teaching can be found in the literature. Hickey and his colleagues analyzed 30 community digital badge design projects, and summarized 26 design principles for digital badges (Ian O'Byrne et al. 2015). Charleer and his colleagues have implemented badge systems in university teaching. They concluded that visualizations can support students in effectively exploring their efforts and outcomes (de Freitas et al. 2017). Some research findings, however, are negative, if the badge system is not well designed. For example, Corrin revealed that some students were distracted by the dashboard from their overall performance goals or were not able to interpret the feedback from the dashboard (Corrin et al. 2015). Based on our review of the literature, we discovered that the development of most badge systems was not usually guided by or evaluated in accordance with evidence-based pedagogical principles. We believe that a principle-oriented framework is needed for guiding and/or evaluating the implementation of badge systems.

The paper evaluates the effectiveness of a badge system in an outreach program based on Chickering's "Seven Principles for Good Practice in Undergraduate Education" (Chickering et al. 1987). The design of the course and the badge system is described in Section 2. Analysis of system usage is presented in Section 3. Evaluations of the effectiveness of the badge system based on the seven principles is then presented in Section 4. HCI-focused evaluation based on Object-Action Interface (OAI) model is described in Section 5. Future development based on the evaluation is presented in Section 6.

2 DESIGN OF THE BADGE SYSTEM

2.1 Course Overview and the Needs of a Badge System

HSST9003 Everyday Computing is a 41-day non-credit-bearing outreach course offered by the University of Hong Kong (HKU) for talented high school students aged 15-19 from all around the world. The cohort studied in this paper was delivered using a flipped classroom/boot-camp approach - students first remotely learned concepts of algorithmic design via online videos and quizzes on Open edX. After learning the basic concepts, they would engage in intensive activities and work on a STEM design project in one-week face-to-face sessions in HKU. Even though the course was not credit-bearing, students could earn a badge by completing one topic in the online course, and would receive a certificate of completion if they earned four out of seven badges. The first five badges were basic badges, and the last two were advanced badges.

HSST9003 was a modified version of a full semester on-campus course for undergraduate students, entitled CCST9003 Everyday Computing and the Internet. The two courses provided feedback to students using two different learning visualization tools – CCST9003 provided feedback using a customized grade-centric learning progress dashboard (Hu et al. 2017), while HSST9003 offered immediate non-score-based feedback with gamified badges. The purpose of introducing the badges was to intrinsically motivate young learners to explore each topic and accomplish tasks paired with it.

2.2 Design of the Badge System

There are two interfaces, one for students, and one for teachers. In the student interface, the number of badges earned is shown on the top bar, as shown in Fig. 1(a). Coloured badges are those earned by students, while monotone ones are yet to be collected. This interface provides an overview of students' online learning progress. Students can earn a badge (i.e., changing a monotone badge into a colored one) by visiting designated courseware pages, watching videos, attempting challenge questions and attaining 50% of the total score in each topic. In addition, by clicking the monotone badge, the badge system indicates what actions are yet to be taken in order to earn the badge, as shown in Fig. 1(a). The interface visualizes students' progress and students can use it as a basis when deciding on their next steps in learning. Meanwhile, the instructor interface (as shown in upper part of Fig. 1(b)) provides a real-time summary of badges earned by a particular student and unfinished videos. The interface also lists out details of certificate earners, as shown in the lower part of Fig. 1(b).

The badge system is implemented as a customized edX XBlock, which is a basic component for building edX courseware. Teachers first specify the course components to be measured through XBlock edit mode. The server of the badge system can then directly fetch data from the database and transfer analyzed results to the client through AJAX. Through visualizations generated with the D3 library, the badge system enables students to check his/her own online learning progress in near real-time. In order to make the badges more attractive to young learners, we worked with a design artist to produce colorful and topic-related badges with "cool experts" names as shown in Figures 1 and 2. Furthermore, titles for corresponding badges are customizable in the system.

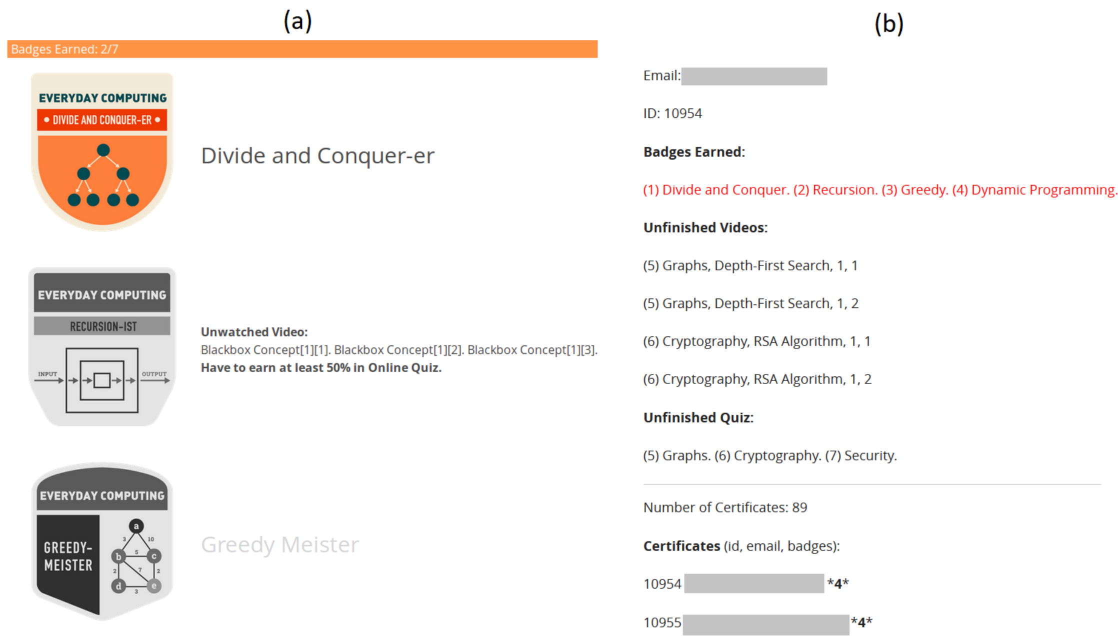


Figure 1: (a) Student interface of the badge system, and (b) Teacher interface of the badge system.

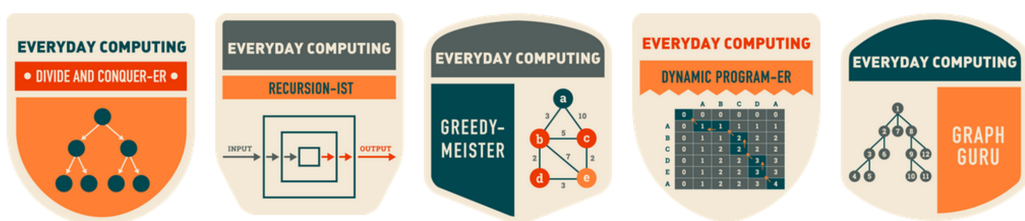


Figure 2: Five colorful basic badges in the badge system.

3 USAGE ANALYSIS OF THE BADGE SYSTEM

All 97 students accessed the dashboard 1153 times within 41 days (19 June - 28 July), and checked the dashboard for about one minute per visit on average. The exit rate, which is the percentage that were the last browsed page in the visit, was only 11.19%, indicating that learners usually used the badge system as a portal for browsing other courseware pages, rather than as the last terminal of the learning journey. Therefore, we claim that most students used the badge system for defining their learning goals in that visit, but not for checking the completion of learning tasks after the visit.

Table 1: Distribution of collected badges (* 0 – 3: No certificate will be rewarded).

Number of badges collected	0 – 3 *	4	5	6	7	Total
Number of students	8	53	18	11	7	97

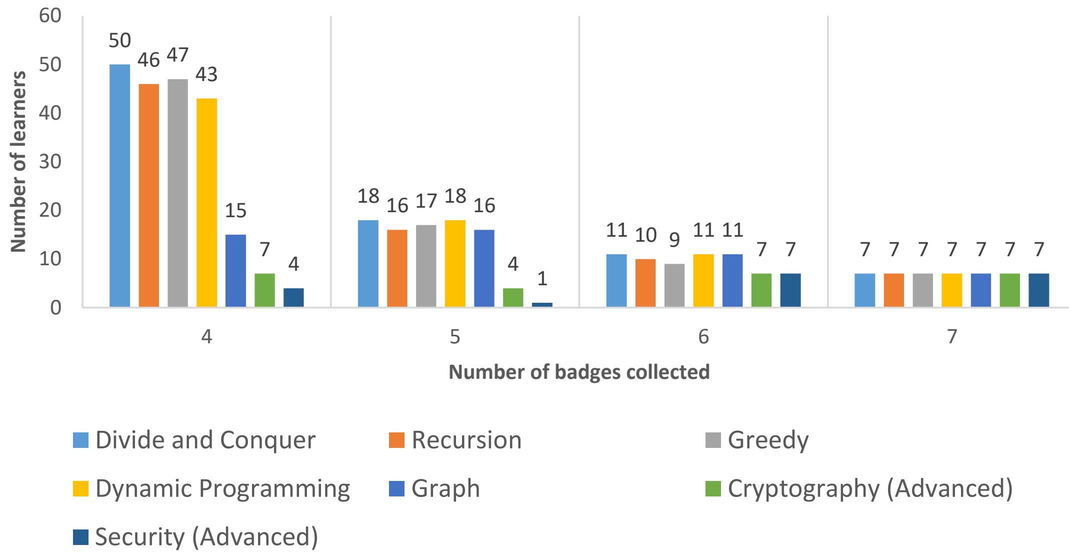


Figure 3: Distributions of badges collected by students who earned more than four badges.

By the end of the course, 89 students received the certificate. Among these students, 13 students did not earn all five basic badges, but obtained one or two advanced badges. Meanwhile, 36 out of 97 students collected more than four badges, exceeding the minimum requirement for getting the certificate. In particular, 7 of them got all badges, as shown in Table 1 and Fig. 3. This indicates that the learners were self-motivated to learn: They were not just aiming for an optional certificate but were also aiming to finish all contents. Therefore, we claim that the badge system fulfilled its functions as intended. Fig. 4 shows the distribution of earned badges according to the topic.

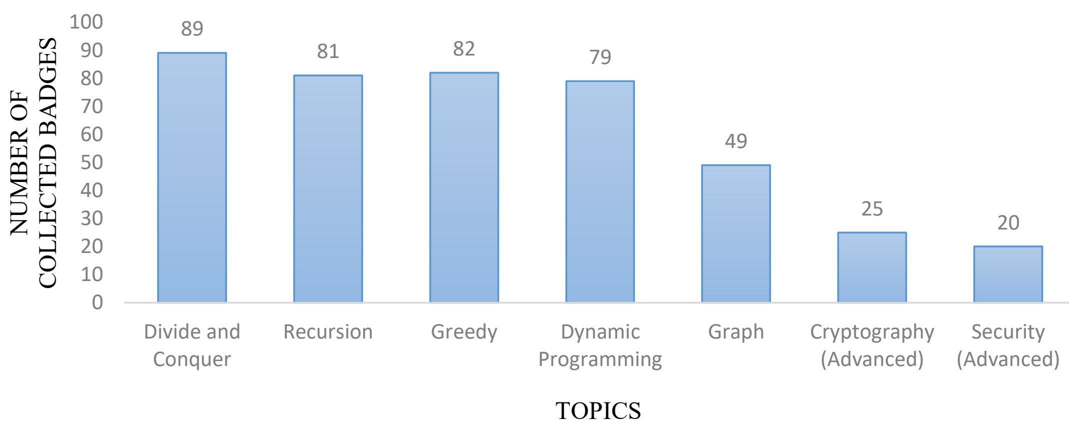


Figure 4: Distribution of earned badges according to the topics in the course.

Fig. 5 shows the daily usage of the badge system during the distant learning sessions (19 June - 28 July). In the soft adoption stage (June 19), not many students clicked the badge system. Students started to check the badge system after the course teacher's first webinar on June 26 where he explained the courseware; and the second webinar on July 6 where he introduced the badge system and the requirement for obtaining the certificate. Most students used the badge system when the deadline of assessments (July 19 and July 20) was approaching.

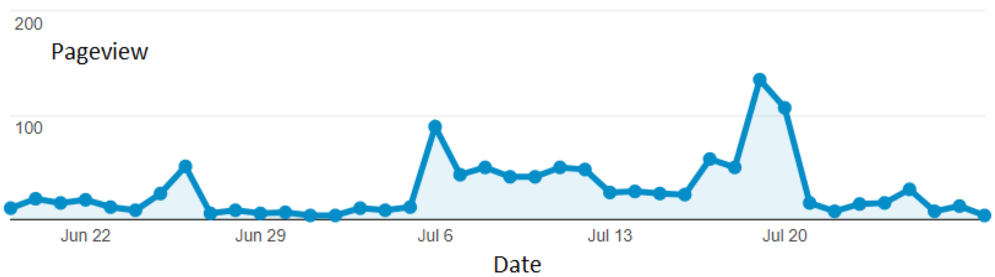


Figure 5: Number of pageviews of the badge system during the online learning sessions

4 EVALUATING THE DESIGN OF THE BADGE SYSTEM BASED ON "SEVEN EFFECTIVE TEACHING PRINCIPLES"

In this study, we evaluated the badge system using the "Seven principles for Good Practice in Undergraduate Education" (Section 4) and Object-Action Interface model (Section 5). The "Seven principles for Good Practice in Undergraduate Education", proposed in 1987, have been widely used for guiding and evaluating design of courses and teaching/learning activities across different educational settings (Bradford et al. 1997; Chickering et al. 1987; Chickering et al. 1996; Chickering et al. 1999; Graham et al. 2001, Luo et al. 2017). Effectiveness of adopting these principles have been proven. Therefore, they will be used for evaluating the design of the badge system.

4.1 Principle 1: Encourages contacts between students and faculty

This principle highlights the importance of having as much (online) student-teacher interactions as possible in order to improve students' motivation and engagement in online learning. Visualizing the learning process (as shown in Fig. 1(a)) using badges enabled students to quickly understand their learning progress in an online context without the teacher's presence. The designed system provided a clear indication of learning quests to be completed for obtaining the badges. This stimulated students to initiate discussions with teachers on the online forum to seek clarifications of learning contents for collecting the badges.

4.2 Principle 2: Develops reciprocity and cooperation among students

This principle aims to minimize competitions among students, and provide a friendly environment for peer-learning. In the designed system, instead of providing feedback using a leaderboard, the badge system effectively summarized students' individual performance without ranking them (as shown in Fig. 1(a)). In other words, every student's achievement was independent of their peers. This encouraged students to help each other in learning and explain to each other how to obtain the badges. This type of conversation frequently occurred on the discussion forum in the studied course.

4.3 Principle 3: Encourages active learning

This principle encourages students to learn through active engagement in activities rather than passively receiving information. In the designed system, instead of measuring a single type of activities for earning badges, students were required to complete various tasks (e.g., browsing specific pages and videos as well as attempting quizzes) in order to obtain a badge (as described in Section 2.2). Furthermore, students could click the interactive badges to retrieve more feedback and identify the actions to be taken next in order to earn the badges (as shown in Fig. 1(a)).

4.4 Principle 4: Gives prompt feedback

This principle is about facilitating students' learning through prompt feedback. The badge system provided real-time recognition of learning achievements, on the top bar in the interface throughout the course (as shown in Fig. 1(a)). This facilitated learning as well as built up students' self-confidence in further exploring the topic. The badge system also listed out videos and materials that were yet to be explored for collecting the badges. This served the important function of guiding students what contents to explore next.

4.5 Principle 5: Emphasizes time on tasks

This principle aims to raise students' awareness to the importance of making good use of their time for learning. In order to manage students' expectations and to guide them through the badge system, the course team set up an administrative page that listed out all the basic information, requirements and mechanisms for obtaining the badge and the certificate, such as deadlines for collecting badges. Furthermore, the same information was delivered to each student via email and was regularly posted on the course update page. In general, students began to ask for clarifications and raised other concerns they had regarding the badges one day after they received the information.

4.6 Principle 6: Communicates high expectations

This principle recommends teachers to explain and communicate their high expectations to learners, so as to motivate learners to strive for better performance. In this course and designed system, we offered five basic badges on algorithm design for all students, and two advanced badges on optional topics (internet security) for high achievers who wanted to learn more. Results in Fig. 3 indicate that the learners were self-motivated to learn: They were not just aiming for an optional certificate but were also aiming to finish all contents.

4.7 Principle 7: Respects diverse talents and ways of learning

This principle encourages teachers to respect students' choices in deciding their own learning paths based on interests. In the studied course, as students were only required to obtain four out of seven badges in exchange for the certificate, they were free to choose how many badges to collect and which/when to collect. It was also their choice as to whether or not they invest the effort to obtain the certificate. Given that part of the learning content was available online, students could work at their own pace and time zone. As the students were from all around the world with different backgrounds, expertises, and interests, this principle was important in catering to students' diverse needs.

5 EVALUATING THE DESIGN OF THE BADGE SYSTEM BASED ON OBJECT-ACTION INTERFACE MODEL

We have asked a human-computer interface expert for comments on system interface design. He agreed that besides providing neat-design badges, learners could easily retrieve hidden informative feedback (i.e., actions that are yet to be taken in order to the badges) by clicking the badges, instead of constantly approaching the teaching team through emails. On the other hand, he proposed that the design of the system can also be evaluated by the Object-Action Interface (OAI) model (Shneiderman 2010), i.e., the badge system should be anchored to student-familiar concepts with a logical structure. For example, existing figures on the badges (e.g., recursion in Figures 1 and 2) cannot be easily comprehended by novice learners. Furthermore, a three-level badge system (e.g., “Recursion-ist trainee”/”Junior recursion-ist”/”Senior recursion-ist”) could be implemented, instead of the current two-layer badge system (Not achieved/Achieved).

6 FUTURE DEVELOPMENT BASED ON THE EVALUATION

In the future cohorts, in addition to content-based badges, we plan to introduce action/behavior-based badges to further encourage active learning behaviours as well as enhance learning interests and motivations. Examples of behaviour-based badges are as follows:

- “Too Active to Have a Rest”: Logged in the online learning platform for 5 consecutive days. (Addressing Principles 1, 5 and 6)
- “Gentle Poster”: Received a reply and a vote in the discussion forum. (Addressing Principles 1, 2 and 3)
- “Endless Patience”: Watched a video from start to end without skipping a second (Addressing Principles 5 and 6).

To further promote active learning (Principle 3), we will also design badges for on-campus activities, in order to capture outcomes and behaviors in the active and engaging face-to-face sessions.

We also aim to generalize the system into an open-source edX Xblock, such that it can be used in other courses. For the purpose of generalization, we will revise the system for querying live database, such that the course structure can be automatically extracted for any specific course. Furthermore, badges should be customizable by course teachers. Technically, we will also revise the teacher interface of the badge system, in order to show the distribution of earned badges in the class.

7 CONCLUSIONS

The paper evaluated the implementation of a badge system based on seven principles of effective teaching. Based on our analysis of system usage, we claim that the system had intrinsically motivated students to participate and pursue achievement in the course. Meanwhile, the system can be further enriched by adopting human-computer interface design principles, for ease of interpretation by students. It is hoped that the findings in this paper can inspire teachers and e-learning technologists to adopt effective badge systems and other learning visualization tools for teaching and learning.

REFERENCES

Auvinen, T., Hakulinen, L., & Malmi, L. (2015). Increasing students' awareness of their behavior in online learning environments with visualizations and achievement badges. *IEEE Transactions on Learning Technologies*, *8*(3), 261-273.

Bradford, B. M., & Peck Jr, M. W. (1997). Achieving AECC outcomes through the seven principles for good practice in undergraduate education. *Journal of Education for Business*, *72*(6), 364-368.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE bulletin*, *3*, 7.

Chickering, A. W., & Ehrmann, S. C. (1996). Implementing the seven principles: Technology as lever. *AAHE bulletin*, *49*, 3-6.

Chickering, A. W., & Gamson, Z. F. (1999). Development and adaptations of the seven principles for good practice in undergraduate education. *New directions for teaching and learning*, *1999*(80), 75-81.

Corrin, L., & de Barba, P. (2015, March). How do students interpret feedback delivered via dashboards?. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 430-431). ACM.

de Freitas, S., Gibson, D., Alvarez, V., Irving, L., Star, K., Charleer, S., & Verbert, K. (2017, April). How to use Gamified Dashboards and Learning Analytics for Providing Immediate Student Feedback and Performance Tracking in Higher Education. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 429-434). International World Wide Web Conferences Steering Committee.

Graham, C., Cagiltay, K., Lim, B. R., Craner, J., & Duffy, T. M. (2001). Seven principles of effective teaching: A practical lens for evaluating online courses. *The technology source*, *30*(5), 50.

Hu, X., Hou, X., Lei, C. U., Yang, C., & Ng, J. (2017, March). An outcome-based dashboard for moodle and Open edX. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 604-605). ACM.

Ian O'Byrne, W., Schenke, K., WillisIII, J. E., & Hickey, D. T. (2015). Digital Badges: Recognizing, Assessing, and Motivating Learners In and Out of School Contexts. *Journal of Adolescent & Adult Literacy*, *58*(6), 451-454.

Luo, C., Hew, T.K.F., Lei, C.U., Oh, E. (2017, December). Principle-Guided Flipped Classroom Implementation Framework for Teaching Technological Contents'', in *Proceedings of the IEEE Intl. Conf. on Teaching, Assessment, and Learning for Engineering* (pp. 449-456). IEEE.

Hu, X., Hou, X., Lei, C. U., Yang, C., & Ng, J. (2017, March). An outcome-based dashboard for moodle and Open edX. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 604-605). ACM.

Mah, D. K. (2016). Learning analytics and digital badges: potential impact on student retention in higher education. *Technology, Knowledge and Learning*, *21*(3), 285-305.

Shneiderman, B. (2010). *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*(6), 1499-1514.

Scaling Nationally: Lessons Learned

Michael Webb
Jisc
michael.webb@jisc.ac.uk

Paul Bailey
Jisc
paul.bailey@jisc.ac.uk

ABSTRACT: A national learning analytics service has been under development in the UK, led by a non-profit organization with universities, colleges and other post sixteen education providers as members. After two years of development the project is moving to full service mode. This paper reports on seven of the key lessons learnt so far from the first twenty pathfinder organizations, along with the transition-to-service process expanding to other organizations. The lessons cover the makeup of the project team, functionality of services, the speed of change processes, the success of standards, legal complexity, the complexity of describing predictive models and the challenge of the innovation chasm. Although these lessons are from the perspective of a service provider, most should be equally applicable to the deployment of analytics solutions within a single organization.

Keywords: service deliver, pilots, national, teams, legal, tools

1 INTRODUCTION

This paper reports on seven of the key lessons learnt from the development of a national learning analytics service. The service has been developed by a national non-profit organization, with around 600 post-16 colleges, universities and other education providers as members, with the aim of accelerating the adoption of learning analytics nationally.

The service has run as a development project for two years, involving around 20 universities and colleges and six technology partners, and is moving to full service mode in quarter four 2017.

The aim of the service is to provide a core infrastructure for learning analytics, including data extraction and storage, a set of applications to allow institutions to start learning analytics pilots, a set of APIs to allow vendors to build upon the data and architecture, and the supporting toolkit and community events required to help make learning analytics initiatives a success.

2 Success Criteria

For the initial two-year phase of the project to have been considered a success, two main criteria needed to be met:

a) The project had to deliver a service that our member institutions wanted. This is measured by the number of institutions continuing beyond the project phase into service mode.

b) The project had to lead to a service that was financially viable. Income from the service needs to cover its costs, with the service delivering a positive net present value over five years.

The lessons noted here focus on the first criteria.

3 Seven key lessons learnt

The project was run using an agile process, with regular retrospectives, as well as regular review meetings with the pathfinder organizations, with many lessons being learnt along the way. This paper brings together seven of the core lessons that may be useful to other organizations attempting to deliver a similar service. The lessons covered, which are expanded on in section 3, are as follows:

- Lesson 1: The team needs a number of core roles in order to succeed
- Lesson 2: The tools should be developed with users and match their terminology and processes.
- Lesson 3: Do not expect process change to occur quickly.
- Lesson 4: Applying standards to data really does work
- Lesson 5: Do not underestimate legal and contractual complexity
- Lesson 6: Users want to understand predictive models (and that is hard)
- Lesson 7: Consider the innovation chasm

4 A brief overview of the service

The service contains two main parts - the technology and the tools and community that institutions need to make the most of the service.

Core elements of the technology include:

- A cloud based multi-tenanted standards-based store for learning analytics data
- Plug-ins, connectors and tools to allow institutions to submit data to the store
- APIs to allow vendors to extract data from the store
- A predictive modeling service

- A staff dashboard and student app

Core elements of community and tools include:

- Networking events held four times a year with around 100 attendees per event.
- A code of practice covering legal and ethical issues
- A mailing list with around 600 members and a blog with over 60 posts
- A procurement framework agreement
- Documentation and guides

5 Seven Key LESSONS

6 Lesson 1: The team needs a number of core roles in order to succeed

During the pilot stage with the first twenty institutions, no specific guidance on the makeup of the project team was prescribed. The only requisite was that there should be demonstrable senior management buy in. Teams were typically lead by either by a learning technology function, or the IT department, and contained a mix of academic staff representation, data owners, technical staff, project managers and teaching and learning specialists.

Process varied significantly between the pilot groups. The teams that made the best progress, defined as moving to live pilots with academic staff had the following characteristics:

- **A single senior manager with clear responsibility for the project**. It does not seem to matter which part of the organization this role comes from.
- **A dedicated project manager.**
- **A named contact for each department/service responsible for delivering data**, and with those contacts having dedicated time to deliver the data.
- **A number of named academic staff representatives.**

The impact of the lack of each of the role is summarized in Table 1.

Table 1: Impact of role not present in project

Roles	Impact if not engaged
Senior Manager	Project never actually starts
Dedicated Project Manager	Slow progress

Named Data Contact	The project does not progress to pilot, as required data is not obtained
Named Academic Staff representatives	Lack of feedback on tools, leading to inappropriate choices Difficult to progress from working prototype to pilots with users No buy-in from academic staff community

7 Lesson 2: The tools should be developed with users and match their terminology and processes.

Initial versions of the tools developed for institutions took a very functional view of learning analytics. For example, once a predictive model had been run, the results would be presented. The language used in the tools reflected the mechanics of the models and visualizations, for example, describing the outcomes as calculations of risk. In early testing this approach was rejected by users.

The core issues were as follows:

- The tool did not obviously fit into users existing workflow and processes.
- The information presented did not give enough information for the staff member to take informed action.
- The language used, particularly around risk, did not match users view. They were interested in success, not risk.

To address this, a new tool was developed, providing dashboards and visualizations fitting around existing roles and processes. The initial roles are:

- Personal Tutors: People who meet with students several times a year, and review progress.
- Module Leaders: People who want to understand how the teaching their module is going.

In addition, a classic agile development approach was adopted, with monthly meetings with user groups determining which features are to be developed next.

These changes have meant that users now understand and buy-in to the tools.

8 Lesson 3: Do not expect process change to occur quickly

One of the original design concepts of the architecture concerned how interventions would occur once a student was discovered to be at risk of failing. The assumption was that an alert would be sent to an individual or team responsible for student success, and they would initiate the required intervention. As the pilot projects progressed, it became clear that most institutions did not have the process in place to deal with alerts in this way. In particular, in all but one of the institutions involved (a small institution with < 3000 students) the team for handling the alert was not in place, and focus was more on refining the personal tutorial systems, as noted in lesson 2.

Whilst it may seem describable desirable to create an additional role, responsible for responding to alerts, in all pilot institutions this would be a significant process and role change, and was considered to be longer term than the initial pilots. Although most institutions had student support services in place already, they had no process for collating the interventions made across academic and central support services.

It was therefore necessary to take a pragmatic approach, and instead use the tools and service to support existing processes and assist institutions to explore further once the initial pilots were completed

9 Lesson 4: Applying standards to data really does work

At the core of the project are two standards, handling activity data and data about the student.

The activity data is collected in xAPI format, with statement templates developed collectively nationally. xAPI was selected as it was the most mature standard at the time of project inception.

There is no suitable international standard for student data, and therefore a national standard was developed - the Universal Data Definitions.

A standards based approach was taken with the aim of enabling models and visualizations to be shared between institutions, regardless of the underlying technology and systems used. This aim has come to fruition, with the following being notable benefits:

- The same visualizations are being used by institutions with different learning management systems and different student record systems.
- Visualizations are being shared across domain spaces, so the same visualizations can be used for library, attendance and learning management activity as the xAPI templates used across systems contain shared core elements.
- Predictive models are being shared across institutions, with different learning management systems, attendance monitoring solutions and different student records systems.

10 Lesson 5: Do not underestimate legal and contractual complexity

During the pilot phase of the project, a three-way legal contract was developed, which covered service provision and data protection issues. In this phase, each part was allowed to request changes to the contract.

This was done in order to make it as simple as possible for institutions to sign up, with the national body leading the project being the flexible body. However, the end result was that the time taken to agree each contract varied significantly, and is some cases introduced much delay. Time to sign ranged from 8 days of days to 183 days, with a mean of 43 days for the first 18 institutions.

The main areas of comment and amendment were as follows:

- Clarity on data fields covered by the contract
- The amount of liability offered in event of data breach
- General drafting issues that do not materially affect the contract.

Whilst this approach provides maximum flexibility, it would not scale, due to the time taken by and with each institution on the legal process. In addition, compliance with the new European Union General Data Protection Regulations (GDPR) became a core area of concerned of all institutions involved.

For the service mode, therefore, non-negotiable contracts were introduced, and all contracts became two-way rather than three-way.

11 Lesson 6: Users want to understand predictive models (and that is hard)

A core concept of the service is that is allows institutions to conform to Jisc's Learning Analytics Code of Practice (Sclater and Bailey, 2015). One aspect covers algorithms:

> "All algorithms and metrics used for predictive analytics or interventions should be understood, validated, reviewed and improved by appropriately qualified staff"

Furthermore, workshops with academic staff showed that they also wished to understand how the models worked.

The underlying assumption from most users was that model was based on rules, and it should show what factors led to a given prediction. The predictive model is actually based on logistic regression and neural networks, and explaining to users from a non-mathematical background how this works is challenging.

The following approach has been adopted:

- A detailed guide aimed on how the model functions has been produced, aimed at relatively numerate institutional staff, to enable them to meet the spirit of the code of practice.
- Academic staff are shown both the prediction, and a number of rule based traffic lights. The traffic lights provide supporting information to help them understand what might be contribution to the success prediction.

12 Lesson 7: Consider the innovation chasm

Moore (1991) presents the concept of the chasm as a gap between the requirements of innovators and early adopters and the rest (early majority, late majority and laggards).

By definition, the institutions taking part in the first phase of the pilots were early adopters, and behaved in a way consistent with Moore's description:

> 'They want to start out with a pilot project, which makes sense because they are 'going where no man has gone before' and you are going with them"

As the project moved beyond the initial pilot group, institutions behaved in a way consistent with Moore's description of early majority.

> "they care about the company they are buying from, the quality of the product they are buying, the infrastructure of supporting products and system interfaces, and the reliability of the service they are going to get"

To address the requirement of the early majority the following additional documentation has so far been required:

- A security guide, detailing the main security processes and features of the service.
- A service level agreement
- A clear pricing structure

13 SUMMARY AND CONCLUSIONS

These lessons are drawn from learning analytics pilots across 20 institutions. Whilst the context of them was from the viewpoint of a provider delivering analytics solutions, the lessons could equally apply within a single institution.

The community element has proven important in sharing experience between institutions, and the process of collating and sharing new lessons is likely to form a core element of the service going forward. It is worth noting that the agile approach lends itself to an action research approach and form of evaluation, so new lessons are likely to be learnt and shared on a regular basis.

Within the project, these lessons are incorporated in an onboarding guide, aimed at helping institutions start their learning analytics projects, along with a strategic guide aimed at senior managers.

REFERENCES

Sclater, N, Bailey,P. (2015). *Code of practice for learning analytics.* Retrieved from https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics

Moore, G. (2014). *Crossing the Chasm, 3rd Edition: Marketing and Selling Disruptive Products to Mainstream Customers* New York: Harper Collins.

Developing an Institutional Policy using the RAPID Outcome Monitoring Approach

Ainsley Hainey
University of Strathclyde
ainsley.hainey@strath.ac.uk

Brian Green
University of Strathclyde
b.green@strath.ac.uk

Helyn Gould
University of Strathclyde
helyn.gould@strath.ac.uk

Howard Ramsay
University of Strathclyde
howard.ramsay@strath.ac.uk

Catherine Milligan
University of Strathclyde
C.Milligan@strath.ac.uk

ABSTRACT: Higher education institutions urgently require policies and strategies for the implementation of the use of learning analytics. Collaborating in research with the Supporting Higher Education to Integrate Learning Analytics (SHEILA) Project, the RAPID Outcome Mapping Approach (ROMA) provided a framework for developing a policy for wide-scale institutional adoption of learning analytics, after an initial small-scale pilot project at the institution. In this paper, ROMA will be discussed, and the resultant mapping for our institution will be used as an example to guide other institutions in their development of policy in this area. Senior management involvement in the process, in addition to the keys to success and lessons learned from this approach, will be discussed in this paper.

Keywords: Learning analytics, policy, higher education

1 INTRODUCTION

At the 7th International Learning Analytics & Knowledge Conference hosted by Simon Fraser University in Vancouver, Canada, we presented a paper on the development of an institutional strategy for implementing the use of learning analytics. We developed this strategy through a Learning Analytics Strategy Group and a Learning Analytics Steering Group. They were guiding and providing institutional oversight to the activities used to capture the evidence, which would inform the resultant strategy. Their responsibilities included:

- Oversight of the scoping of current data sources and analysis tools that existed within the University;
- researching the learning analytics field to identify examples of effective use of learning analytics in the sector;
- designating data into three 'data levels' and formalising how each data level is reported and used within the Institution;
- participating in a Learning Analytics Readiness Assessment undertaken by Blackboard on behalf of Jisc;
- identifying five pilot classes, one from each of the Institution's four faculties in addition to a class provided by the University's Organisational and Staff Development Unit (OSDU), to implement a learning analytics approach providing proof of concept evidence to inform the finalised learning analytics strategy; and
- mapping learning analytics' institutional potential to support, improve and provide evidence for key objectives identified in the strategic documents produced by the University. Included in these documents were the University's Strategic Plan 2015-2020, the institutional Scottish Funding Council Outcome Agreement, the Quality Assurance Agency for Scotland's Enhancement Led Institutional Review and the Scottish Enhancement Framework's emerging sector enhancement theme, 'Evidence for Enhancement: Improving the Student Experience', to name a few.

An additional key responsibility of the Groups was:

- networking with other Institutions and agencies currently implementing or scoping learning analytics, such as the UK Open University and Jisc.

It was through this networking remit, that we participated in the Supporting Higher Education to Integrate Learning Analytics (SHEILA) Project[1] in August 2016. This was a research study exploring how institutions were implementing learning analytics to inform the creation of a policy development framework to support higher education. We followed the project dissemination activities and found that many institutions appeared to be in a similar position to ourselves, in that they had found a lack of policies or guidance for implementing learning analytics, and had implemented small-scale pilot studies (Tsai & Gasevic, 2017).

In addition to presenting at the 7th International Learning Analytics & Knowledge Conference, we attended a workshop at the same conference, delivered by the SHEILA project team - LA Policy: Developing an Institutional Policy for Learning Analytics using the RAPID Outcome Mapping Approach. This gave us an insight into a potential framework approach through which institutions could develop policies for using learning analytics, and without knowing of ROMA previously, we found that it

[1] http://sheilaproject.eu/

reflected key elements of our work and that we had already mapped aspects of the Approach in the work of our pilot projects.

2 RAPID OUTCOME MAPPING APPROACH (ROMA)

The RAPID Outcome Mapping Approach (ROMA) was developed as a tool by the Overseas Development Institute to develop strategies for evidence-based policy-making (Young & Mendizabal, 2009). The Approach provides 8 steps for successfully implementing a policy, however this has been modified to the Approach outlined in Figure 1 below by the SHEILA Project ("SHIELA – Project Approach", n.d.).

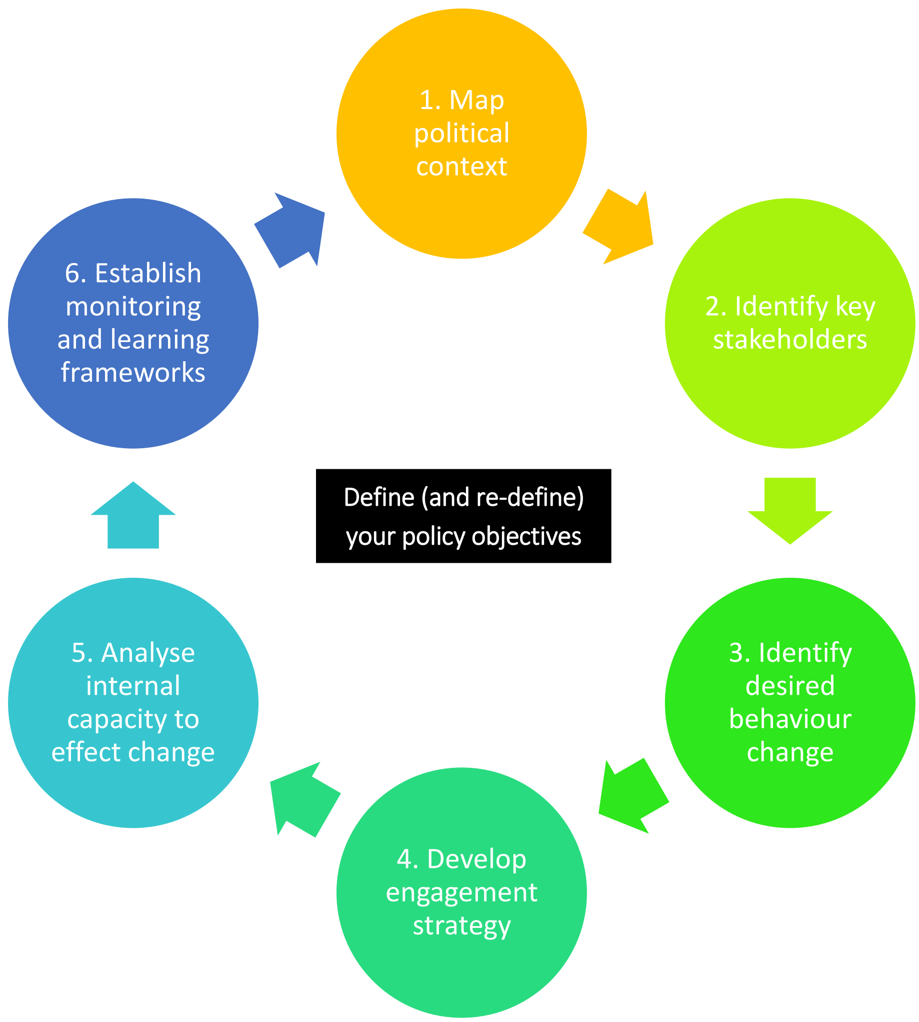


Figure 1: RAPID Outcome Mapping Approach (ROMA)

3 IMPLEMENTING ROMA

From our initial engagement in the SHEILA Project, researchers had mapped our pilot projects against the ROMA framework for us, as this formed part of their case study and they were able to share this documentation with us. Further discussions with them focused on institution-wide adoption of learning analytics and the associated challenges and potential solutions shared by other institutions in their process of adoption, which was invaluable in providing insight into how to map the next phase of our project against ROMA, providing a policy framework for adoption that provides the necessary focus and works for our institution.

Involvement of University senior management and reporting through strategic education committees and groups has been vital in this Approach, particularly at the stage of institution-wide adoption. They have oversight of all other strategic projects occurring, or incoming to the University, which is helpful particularly in the ROMA steps of mapping the political context, identifying key stakeholders, and analysing internal capacity.

Due to our involvement with the SHEILA Project, we have now become an Associate Partner of the Project, which aims to facilitate experience sharing among higher education institutions in the development of institutional capacity for learning analytics.

4 CASE STUDY

To further the understanding and provide a framework for intuitions challenged by the adoption and implementation of learning analytics, we have provided our own ROMA approach as a case study. This is provided in table 1 below.

Table 1: ROMA Approach

ROMA Stage	Definition	Institutional Direction/Method
Define policy objectives	Define objectives/motivations for learning analytics	• Enhance student learning experience
Map political context	Identify internal and external drivers	• Improve UK National Student Survey (NSS) results • Improve assessment & feedback and provide evidence for assessment & feedback policy • No retention issue at present however, there is recognition that expansion in distance learning and work based learning programmes could present a challenge • Institutional decision making
Identify key stakeholders	Identify users of learning analytics	• Students • Academics • Head of Departments/Director of Teaching • Vice Dean Academic/Faculty Dean • Senior Management Team • Professional Services • Institutional Education Committees
Identify desired behavior changes	Identify desired changes for key stakeholders in the current context	• Improved assessment experience and improve quality and timeliness of feedback for students • Provide academic staff with a mechanism to review their own assessment & feedback strategy

		• Provide evidence for course design and course review • Provide measurable data and evidence of success for senior departmental and faculty staff • Identify training opportunities for staff in assessment & feedback area • Improved student survey results and improved UK NSS scores in assessment & feedback
Develop engagement strategy	Scope areas related to ethics & privacy, financial & human resources, internal & external support, methodology, and stakeholder engagement	• Consult relevant policies and code of practice • Establish a Learning Analytics Board, with representatives from key stakeholders • Align learning analytics with other educational strategies, such as Learning Enhancement Framework • Conduct faculty and professional services collaboration sessions with staff and similarly with students to ensure positive engagement based on Agile Methodology • Continue with external engagement i.e. SoLAR, LAK, Jisc, SHEILA Project
Analyse internal capacity to effect change	Evaluate culture, legal frameworks, financial capacity, human capacity, and technological infrastructure	• Jisc Learning Analytics Readiness Assessment provided feedback on culture, processes, people and technology • Work with Information Governance Unit to ensure compliance with incoming General Data Protection Regulations • Creation of a data mart within the institutional data warehouse • Examine internal resource capabilities and seek to fund new appointments if needed
Establish monitoring and learning networks	Establish qualitative and quantitative measures of success	• Improved student satisfaction in student surveys • Increased student attainment • Operational efficiencies and satisfaction for academic staff • Increased NSS scores in assessment & feedback areas • Successful implementation of reviewed/revised assessment & feedback policy and associated staff training

We will also present lessons we have learned from our initial pilot and scoping activity, and the impact that the ROMA approach has had on refining and shaping current projects and the development of our institution-wide Learning Analytics strategy.

REFERENCES

Tsai, Y. & Gasevic, D. (2017). The State of Learning Analytics in Europe. Retrieved from http://sheilaproject.eu/wp-content/uploads/2017/04/The-state-of-learning-analytics-in-Europe.pdf

Young, J. and Mendizabal, E. (2009). Helping researchers become policy entrepreneurs - how to develop engagement strategies for evidence-based policy-making. Overseas Development Institute.

SHEILA – Project approach. (n.d.). Retrieved from http://sheilaproject.eu/project-approach/

The Learning Analytics Portal: The Development of a National Community Resource for Learning Analytics

Jo Dugstad Wake
Uni Research Health
jo.wake@uni.no

Nina Morlandstø, Khoa Pham, Barbara Wasson
SLATE, University of Bergen
nina.morlandsto; khoa.pham; barbara.wasson@uib.no

ABSTRACT: This paper describes the development of “Læringsanalyseportalen” (English: The Learning Analytics Portal) through the Map LA project. In 2016 the Norwegian Ministry of Education funded a national centre for learning analytics, SLATE, which opened in June 2016. Part of SLATE’s mandate is to undertake a national survey the state of learning analytics, including limitations and opportunities. To serve this purpose the Map LA project created an online community tool inspired by wikis, web 2.0 and dynamic mind maps, for the Norwegian learning analytics community to be able to continually map and keep track of itself. The tool was developed in close dialogue with representatives from the diverse stakeholders involved in the learning analytics community.

Keywords: LA community tool, dynamic mind map, wiki, learning analytics portal

1 INTRODUCTION

The Norwegian word “dugnad” originates from the Old Norse phrase “dugnaðr”, which translates as “help, good deed, force”, which again stems from Old Norse “duge”, and means both “skill, ability” and “virtue”. The concept translates to several languages and cultures, and the general meaning is voluntary, orchestrated work for the benefit of a community, where there is no reward for participation and no penalty for non-participation. In more recent times it is mostly associated with carrying out practical/fundraising work in a context where there is some kind of social connection or community between the participants, for example membership in sports clubs, non-profit organisations or housing cooperatives. The SLATE Centre (Centre for the Science of Learning & Technology) at the University of Bergen, founded in 2016, has as part of its mandate to survey and provide an inventory of the field and community of learning analytics (LA) in Norway. The Map LA project was initiated to meet this end, and it was decided to carry out this task within the concept and vision of a *national dugnad*. To mediate the creation of the inventory for the LA community “Læringsanalyseportalen”, or “The Learning Analytics Portal” (LAP hereafter) was developed. This paper covers the development, structure, and functionality of LAP.

The overall orientation in the Map LA project is “to identify and support a national community of stakeholders involved or interested in learning analytics”. As such this paper and project speaks to the LAK topics of innovative new tools and techniques, collaboration and sharing, and solving new problems.

LAP is a recently developed digital service provided for the Norwegian LA community to map itself. It is inspired by wikis, web 2.0, and digital mind maps. It covers all material aspects of LA, such as organisations, people, publications and dissemination, projects and activities, applications, and educational data sources. The aim is to connect the community members with each other, and with information about what is going on in the community, for example for a teacher who is interested in LA with other teachers who have participated in projects or interesting applications, or for researchers to find new publications on LA. The goal is support the formation of LA community, and to foster collaboration. Previously members from the different sectors of the LA community had their separate meeting arenas, whilst the goal for LAP is to provide an arena for the whole community.

The LAP initiative reflects that learning analytics is a field that particularly gains from openness and common standards, as seen for example in the benefits of establishing a shared vocabulary for educational data, and that what can be achieved through cooperation is much more than if everyone works in isolation. This paper presents the Map LA project and the development process and functionality of LAP.

1.1 The Map LA project

The Map LA project, initiated by SLATE, has two main goals. One is to identify main stakeholders within academia, industry, the educational sector and government, and uncover their visions for and critical reflections over learning analytics in Norway. The second is to provide an inventory over people, organizations, activities, technology, publications and data sources.

1.2 Related work

Discussing the challenges related to going from "successful TEL prototype" to a tool being taken up and used in educational practices, Scanlon et al. (2013) portrayed the TEL complex as a highly interconnected and many-faceted, with a diverse set of actors, agendas and contexts. The Norwegian learning analytics community can be seen as an instance of the TEL complex, with actors from ICT industry, academic research on pedagogy and technology, students and educators in all levels of education and so on. LAP is designed to be a community building service for this community. Several authors (Suchman, 2007; Nardi & Engeström, 1999) argue that it is a central characteristic of the post industrial/ knowledge economy that the outcomes of work are increasingly disembodied and invisible. In this context, according to Leonardi (2014), workers are engaged in work in "sitting at computers typing reports, performing analyses, writing copy, and performing other tasks that are difficult for observers to discern" (p. 796). We see LAP as a tool for countering developments such as these.

LAP is a tool for facilitating engagement in a community, inspired by wikis and highly visual depictions of information and relationships. LAP is basically about using social media to foster community awareness and cohesion. Tredinnick (2006) defines social networking sites as sites driven by user participation and user-generated content. Flouch and Harris (2010) have studied participatory neighbourhood websites/Web 2.0 for their potential in mobilizing social engagement, participation and cohesion, and found that these websites increase connections between residents, likelihood of active contributions to neighbourhood improvement, and also the sense of belonging to the community. Lovejoy and Saxton (2012) have found benefits with social media platforms for non-profit organizations to engage with stakeholders, which are unavailable with non-interactive websites, namely that they are built for interactivity and communication. Analysing the tweets of 14 non-profit organisations they found that they use Twitter to (1) *inform* the public, (2) *communicate* and engage with the public through community-forming dialogue, and (3) to call for *action.* An important difference between the goals for LAP and other non-profit organisation communication needs is that LAP is more neutral; the main goal is to connect the stakeholders rather than to disseminate a particular message.

2 THE LEARNING ANALYTICS PORTAL

The following section presents the architecture, contents and functionality of LAP.

2.1 Architecture

LAP is constructed as a 3-layered structure, where both the data and business logic are stored as an Oracle database, and the presentation layer is managed by an Apache web server and accessed through a web browser (Figure 1). The end user requires only a supported web browser to use LAP, and latest version of Microsoft Edge, Mozilla Firefox, Safari, and Google Chrome are supported.

LAP comprises a development environment and a production environment, run by the IT-department at the University of Bergen, to support testing of new functionality before it is implemented. The business layer is developed using Oracle Application Express (APEX). The presentation layer is developed using the d3.js (https://d3js.org/) visualisation library, in addition to selected JavaScript libraries.

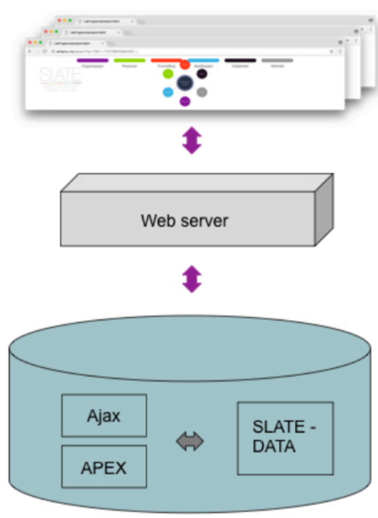


Figure 1: LAP architecture

2.2 Use and users

There are three authorisation levels on LAP -- unauthorised access, authorised access and moderator access. The security concept resembles that of other wikis, such as Wikipedia, where information is available without logging in. To alter or add information, a user has to register and sign in. Creating an account is open to anyone, and based on trust in public use of LAP. Any information is available for view without authorisation, whilst updating or adding information requires a user to first login. Moderator accounts have been created to sort any problems for other users, and survey the information in LAP.

2.2.1 User interface: Viewing information

The most obvious way of using LAP is to peruse it for information about LA. There are three main information structures available for viewing in LAP: tree diagrams, lists and charts. The original idea for LAP was a community-editable, navigable online mind map/tree diagram structure (Figure 2), and this was later augmented first with lists, and then with charts. The goal was to make visible the different aspects of the Norwegian learning analytics community, and important relations between the different aspects. The community includes public and private academic/research institutions, teaching institutions on all three levels, industry, NGOs, publishers, public service organisations, and government bodies on national, regional and local levels.

Six main categories of information were identified, and are visible in the tree structure. Each category has a persistent colour, to support understanding of how kinds of information belong together when navigating the tree structure. Navigation is through clicking the nodes; clicking a node for a category opens the associated subcategories, and further clicking these takes the user to the nodes with associated information. Clicking on the same node once again, takes the user back to the previous view. Each node displays a title for the information it contains. It is also possible to navigate the tree by using the legend presented in the centre top of the screen (Figure 2). Clicking any end node containing information opens a small circle with a context-dependent summary of the information in the bottom left corner. Clicking this circle expands it, and more detailed information is presented, if available.

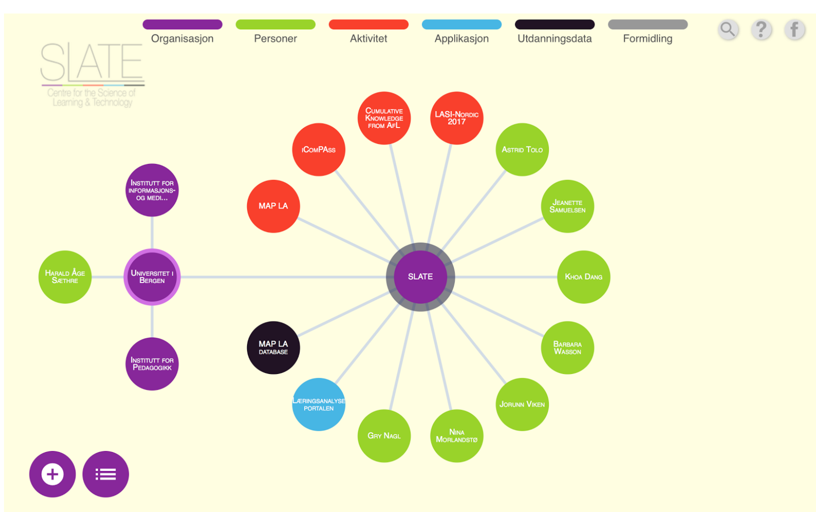


Figure 2: The LAP tree diagram

When a user first enters the LAP interface, a black circle is visible in the centre of the screen. Clicking the circle opens the tree diagram, with the six main categories of information: organisation, activities, persons, dissemination, educational data sources and data applications. Clicking on a main category removes the other categories from view, and presents the subcategories within each category. The complete list of categories and their subcategories is presented in Table 1. Clicking on the node for the subcategory reveals the information that is available within the category.

Table 1: Categories and subcategories in LAP

Information type	Subcategory
Person	Primary school, secondary school, tertiary education institution, private company, government organisation, publisher, research institution, network organisation, other (person)
Organisation	Primary school, secondary school, tertiary education institution, private company, government organisation, publisher, research institution, network organisation, other (organisation)
Educational data source	Statistical data, log data, assessment result, activity data, administrative data, survey data, health data, portfolio data, student survey data, other (educational data source)
Activity	Workshop, meeting, education, project, conference, other (activity)
Application	Statistical, pedagogical, analysis service, API, student survey application, infrastructure, administration, assessment, other (application)
Dissemination	Information, research result, government white paper, media news item, presentation, master/PhD thesis, other (dissemination)

Clicking the list icon (second circle from left in the bottom left corner) displays the information within a subcategory in the form of a list, which offers more information in a more condensed space, more useful for subcategories where there is a lot of information available. The lists can be adapted by each user, and can be manipulated by adding sorting and ordering filters, and adding and removing columns depending on a user's preferences.

The most recent addition to the portal within presenting information is graphs, where summaries of information are presented in a graphical form. Clicking the context-dependent pie chart icon in the bottom left corner, third from left, opens the graphs. The available charts are basic at the time of writing this paper, and will be further refined as ideas for them come to the fore.

2.2.2 User interface: adding, modifying and updating information

Clicking the icon with the "+" sign allows a user to log in to add or modify information on LAP in the editing interface (Figure 3). The editor part of LAP allows a user to create new, delete or edit existing information by the category, and also to make connections between the kinds of information where relevant (tying a publication to a person, or an application to an organisation, for example). Users can carry out actions tied to their account in the topmost right corner of the screen, and navigate directly to the tree diagram view in the top left corner. There is a window for searching for existing information to the right of the screen, and a list of the categories on the left. The left column can also be used to displaying information as lists by the category. The centre of the screen is used to work with information; adding, deleting modifying and making connections.

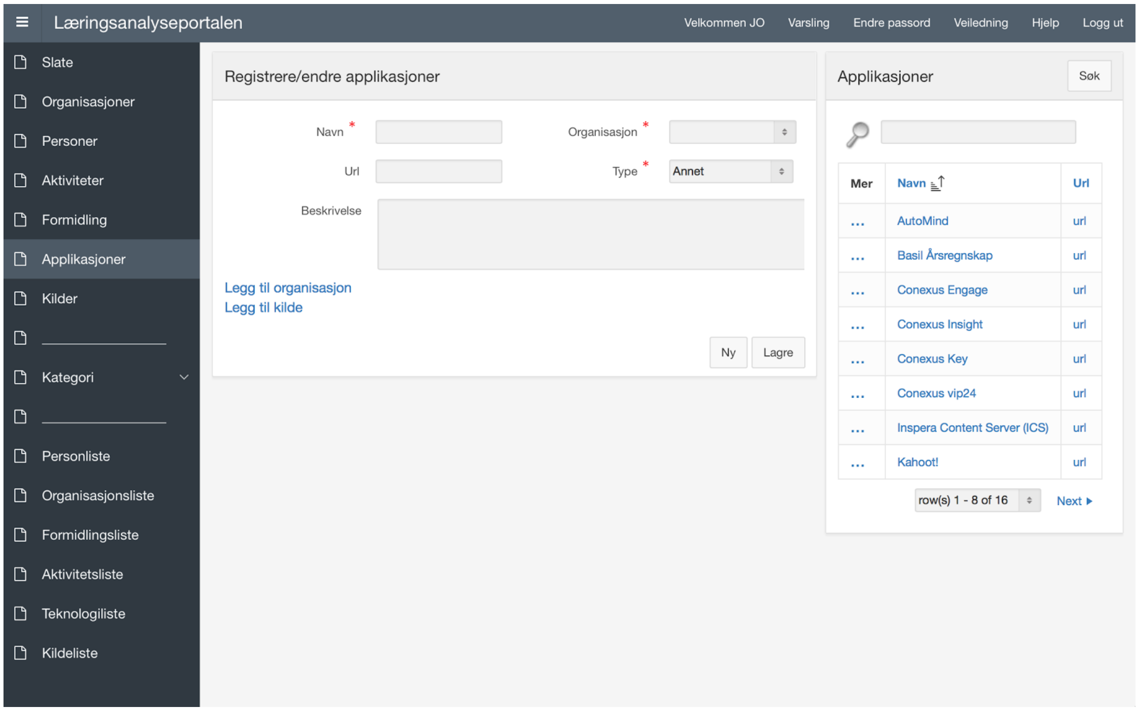


Figure 3: The LAP editor

2.2.3 *Administrative view*

The administrative view within the editor allows *administrators* to create new categories and add definitions to them. Administrators are also able to trace most activity within the editor and to help users with lost passwords, etc. *Moderators* are able to change the categories, delete information, give new passwords to other users, upload files and create new moderator users. Moderators can also generate lists, such as new items in the subcategory “other” (all subcategories have an “other” item, to avoid the loss of new information that doesn’t fit the model), lists of users/people names, lists of changes to the model, and so on.

3 THE DEVELOPMENT PROCESS AND MOVING FORWARD

The development and refinement process has occurred continually since the first version in 2016, and in dialogue with representatives from the learning analytics community. Central facets of the LA community in Norway are the wide variety of roles and geographical dispersion. Representatives include academics from several institutions, industry, publishers and governmental institution employees, including local government. Recruitment for beta testing was carried out through face-to-face requests at a meeting series organised by Standards Norway on education standards in Norway, and by email to scholars who have contributed to the LA discourse in Norway. The participation rate was high, although some organisations where several participants were recruited, delegated participation to one representative from their organisation. To keep track of the feedback and corresponding action steps, comments were gathered in a shared feedback log along with screenshots/explanations and interpretations. Most steps of the development and refinement process have been minor and incremental in nature, and too fine grained to report here.

The kinds of feedback were quite diverse, and the different testers responded to different aspects of LAP. As a whole, the development process has been characterized by mending several smaller bugs and problems. Although the initial idea was to provide a minimalist, dynamic interface with little disturbance in order to highlight the connections between the different pieces of information, several respondents asked for a more structured and clear way of navigating the information in LAP. This included the opportunity to quickly switch between the editor and information views, and more different ways of navigating up and down the nodes of the tree diagram. To meet these responses, we added the legend, visible in the top of the screen in figure 2, enabling access to a category from anywhere in depth in the search tree, and also a button for jumping to the tree diagram from the “edit information” view. To accommodate these views we also added the context dependent icons in the bottom left corner (1 add information, 2 view list and 3 view chart (where charts are available)). The addition of viewing information as lists was also in response to this feedback.

The response from a usability expert among the participants was to increase text size, use bold text in the nodes, and to tweak the colour scheme (originally different). The text size/boldness was amended, and we experimented with the colour scheme before settling on the current background.

Other feedback from testers was about naming of categories, which was also discussed extensively within the developing team. For the developers the challenge has been to use category names that are general yet descriptive enough, and preferably can fit in the node bubbles without breaking the text. For example, “Educational data” was originally called just “sources”, which caused some confusion among the test users and was changed to Educational data sources. Another classification challenge was between the naming of “infrastructure” and “API”, so in the end both were included.

Our own concern is to do with the accumulation of information and its granularity. Presenting information in a tree diagram can reach a point where there is too much information to be presented as nodes in a tree structure, making it difficult for a user to see any particular information at all. This observation has led the project team to consider introducing further methods of categorising the information from subcategories (for example news items and data sources).

Another feedback pointed to the importance of creating awareness about LAP in the broader educational community. Our strategy to raise and sustain awareness was to create a LAP Facebook page. This page also provides the opportunity to engage in dialogue directly with users. We also set up Google Analytics for LAP to learn and know more about the use patterns. Finally, we set up searches for LA in national media archives, to be able to provide LA news on a regular basis.

4 CONCLUSIONS

LAP was created as part of a mandate to gain and maintain an overview of learning analytics use in Norway. In contrast to the LACE Evidence Hub, we did not set out to find and record evidence about learning analytics. Rather our goal was to give national stakeholders an easy to use tool that provides an overview of the national scene within learning analytics, both academic and in practice, research prototypes and Edtech tools, academic publications and reports, links to data sets for use in analytics, etc. The response has been mostly positive, and we are open to requests for improvements. Although LAP is a free service for the benefit of the community, we appreciate that it will take editorial work to keep it updated and relevant and we provide that service, and also a fair bit of marketing to bring it to the attention of relevant users. Finally, there has been interest from other countries to use LAP, and from other organisations in Norway that see it as a potential way to share knowledge internally.

Since the launch of LAP in June 2017, 81 persons, 72 organisations, 81 data sources, 17 applications, 68 publications have been registered.

REFERENCES

Flouch, H., & Harris, K. (2010) The Online Neighbourhood Networks Study: A Study of the Social Impact of Citizen-Run Online Neighbourhood Networks and the Implications for Local Authorities. London: Networked Neighbourhoods Group. Retrieved August 28, 2017, from: http://networkedneighbourhoods.com/?page_id=409

Leonardi. P. (2014). Social Media, Knowledge Sharing, and Innovation: Toward a Theory of Communication Visibility. *Information Systems Research 25*(4), 796-816.

Lovejoy, K., & Saxton, G. (2012). Information, Community, and Action: How Nonprofit Organisations use Social Media. *Journal of Computer-Mediated Communication, 17*(3), 337-353.

Nardi, B.A., Engeström, Y. (1999). A web on the wind: The structure of invisible work. Computer *Supported Cooperative Work 8*(1), 1–8.

Scanlon, E., Sharples, M., Fenton-O'Creevy, M., Fleck, J., Cooban, C., Ferguson, R., Cross, S., and Waterhouse, P. (2013). Beyond prototypes: Enabling innovation in technology-enhanced learning. Open University, Milton Keynes.

Suchman, L. (2007). *Human–Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press, New York.

Tredinnick, L. (2006). Web 2.0 and business: A pointer to the intranets of the future. *Business Information Review, 23*(4), 228–234.

Waters, R., Burnett, E., Lamm, A., & Lucas, J. (2009). Engaging stakeholders through social networking: How non-profit organizations are using Facebook. *Public Relations Review, 35*(2), 102-106.

Meta-Predictive Retention Risk Modeling

Risk Model Readiness Assessment at Scale with X-Ray Learning Analytics

Aleksander Dietrichson
Blackboard Inc
dietrichson@me.com

Diego Forteza
Blackboard Inc
Diego.Forteza@blackboard.com

ABSTRACT: Deploying X-Ray Learning Analytics (Blackboard Inc, 2015) at scale presented the challenge of deploying customized retention risk models to a host of new clients. Prior findings made the researchers believe that it was necessary to create customized risk models for each institution, but this was a challenge to do with the limited resources at their disposal. It quickly became clear that usage patterns detected in the Learning Management System (LMS) were predictive of the later success of the risk model deployments. This paper describes how a meta-predictive model to assess clients' readiness for a retention risk model deployment was developed. The application of this model avoids deployment where not appropriate. It is also shown how significance tests applied to density distributions can be used in order to automate this assessment. A case study is presented with data from two current clients to demonstrate the methodology.

Keywords: Retention Risk Modeling, LMS Data, Scalability, Automatization

1. INTRODUCTION

X-Ray Learning Analytics (Blackboard Inc, 2015) is a learning analytics package offered as an add-on to Moodlerooms[1]' 1 clients as well as to institutions that use Moodle on a self-hosted. Among X-Ray's features is a retention risk report usually based in its entirety on data endogenic to the Learning Management System (LMS). The retention risk is assessed with statistical models which are trained and fitted for each institution individually. Back-testing (Dietrichson, 2016, and Forteza, 2016) have shown accuracies in the 90s and have typically been along the lines of the researchers' expectations during modeling. Other cases have been less fortunate –in several cases the recommendation has been not to deploy a risk model at all, since its utility would likely be insignificant or even counter-productive. These cases quickly became a source of some embarrassment since the analytics team was already in meetings with clients at this point in the process. Consequently it became apparent there was a need for a procedure to assess readiness prior to engaging with the client, a model to predict performance of the risk models, in other words: a meta-predictive model.

1 Moodlerooooms is Blackboards managed open-source offering.

2 METHODOLOGICAL BASES

This research is based on some notions that have emerged from prior experience, both in the form of formal research and by ad-hoc observation. This section briefly describes some of the concepts that guided our efforts.

2.1 Customized Models

Multiple research studies on individual courses have found a significant relationship between frequency of use of the LMS and student grades (Fritz, 2011; Macfadyen & Dawson, 2010; McWilliam, Dawson, & Pei-Ling Tan, 2008; Morris, Finnegan, & Wu, 2005; Rafaeli & Ravid, 1997; Ryabov, 2012; Whitmer, Fernandes, & Allen, 2012). The value of LMS data has been far more important than what is found in conventional demographic or academic experience variables in explaining variation in course grades. However, when analysis is expanded to all courses at an institution, several studies have found no relationship or an extremely weak relationship (Campbell, 2007; Lauria, 2015). These findings were in line with the researchers' experience, and congruent with the view that risk models not only need to be customized on a per-institution basis, and also that a likely outcome of a thorough modeling exercise is the deployment valid for only a subset of courses and even several different models for distinct and distinguishable groups of courses.

2.2 Customized Models

Previous work (Forteza & Nuñez, 2016) on course archetypes demonstrated that online courses can be classified into five categories:

1. Supplemental – high in content but with very little student interaction
2. Complementary – used primarily for one-way teacher-student communication
3. Social – high peer-to peer interaction through discussion boards
4. Evaluative – heavy use of assessments to facilitate content mastery
5. Holistic – high LMS activity with a balances use of assessments, content, and discussion

While it may be immediately intuitive that developing a single model (or even model template) to cover these five use-cases, and that use cases (1) and (2) will likely always result in non-performant models, we still wanted to operationalize this distinction and its implication through empirical evaluation. It is also clear that these categories represent a multi-dimensional continuum, and that the named categories refer to the centroids of each cluster. As such there is clearly going to be some overlap and modeling may be possible for courses that straddle one of more of these categories. Real life experience has also indicated that each institution comes with a unique mix of these archetypes as well as other characteristics relevant to the modeling effort.

2.3 Risk Model Performance

The term Model Performance is used loosely to refer to the potential usefulness of a model, rather than as a weighted (or not) proportion of model precision or recall. While several algorithms –for example: Lopez-Raton, Rodriguez-Alvarez, Cadarso-Suarez, & Gude-Sampedro (2014)– exist for optimizing this relationship. The exact balance point will to a large degree depend on each client's needs: the degree to which interventions are planned as a result of predictions made by X-Ray, the cost of those interventions, institutional policy and practical considerations regarding each institutions' ability to act on the information generated by the system.

2.4 Population Parameters

The outcome variable, i.e. that which we are trying to predict, is typically a dichotomized course pass/fail, although cases with qualified pass are also encountered. In either case, a successful modeling exercise necessitates some variance in this variable. This fact allows us to immediately discard institutions with extremely high or extremely low passing rates. For example, an institution which graduates 95% of its students is not a candidate for risk modeling: Simply predicting success for all students would already result in a .95 precision rate. We thus only consider institutions whose population parameters fall within a certain heuristically defined range.

3 RISK MODEL READINESS ASSESSMENT

In order to determine the likelihood of a successful modeling exercise some global course-level measures are considered. These include: passing rate, proportion of students who have accessed the course (in the LMS), number of graded items, number of quizzes, number of assignments, correlation between quiz grades and final grades, correlation between assignment grades and final grades, mean number of access-log entries (clicks) per student and correlation between clicks and final grades. These measures are constructed based on the historical LMS activity. Final grades refer to the course-grades in the LMS or, if the institution does not use the course-level evaluation in the LMS, from an external source, typically the institutional SIS. When substantial use of discussion fora is detected, linguistic variables are also extracted and included.

The courses into are then divided into three categories a) courses that can be used for training a model, b) courses to which the trained model would be applicable and c) discarded courses. The criteria for the second category (b) is somewhat softer than the training data. This gives us an initial estimate of whether we have enough data to train a risk model (a), and an estimate of the proportion of courses in which we would be able to deploy a performant risk model for the client in question. In order for a course to be useful as part of the training set is needs to have have relevant activity, and this activity must be related to the outcome variable (pass/fail or final grade), i.e. it must have discriminatory value. Courses where this is clearly not the case are immediately removed from consideration. For example, courses in which the proportion of passing students is greater that the proportion of students who have accessed the course are not considered, because it is clear that students' access to the course is not relevant for determining the outcome variable.

Additional restrictions are applied and courses further filtered. The filters applied to the training and application categories are summarized in Table 1.

All correlations for these filters are calculated using the point biserial correlation (Glass & Hopkins, 1995) since the outcome variable has been dichotomized into pass/fall.

Table 1: Summary of Restrictions and Filters.

Filter	Training	Aplication
Proportion of students access	> passing rate	> passing rate
Pass-rate	Not 0 and not 1	
Quizzes or Assignments	> 5	> 5
Graded Items	> 10	> 5
Correlation between Assignment and Final Grades	> 0.5	> 0.5
Correlation between Quiz and Final Grades	> 0.5	> 0.5
Clicks per Student	> 500	> 100
Correlation between Clicks and Final Grades	> 0.5	> 0.25

3.1 Reference Institution

The measures found in Table 1 were also been calculated for institutions where a successful modeling exercise had already taken place. These measures were collapsed into a reference institution and used for comparison with candidate institutions. The process is best described by means of an example, or case-study, presented in the next section.

3.2 Case Study

In this section we present anonymized data from two real candidate institutions, both North American Higher Education Institutions. In the following we will refer to them as Candidate I and Candidate II.

Let us first consider the two types of graded items that have shown to be of most importance for predicting the outcome variable in our reference data, namely: grades on quizzes and grades on assignments. Table 2 and Table 3 show the proportion of these two types of graded items for the two candidate institutions as well as the reference.

Table 2: Proportion of Courses with Assignments at Different Levels.

	Candidate I	Candidate II	Reference
> 1	76 %	13 %	90 %
> 5	61 %	7 %	71 %
> 10	32 %	3 %	55 %

Table 3: Proportion of Quizzes in Courses at Different Levels.

	Candidate I	Candidate II	Reference
> 1	88 %	8 %	78 %
> 5	54 %	4 %	50 %
> 10	38 %	2 %	23 %

We see that Candidate I has a solid performance on these metrics, in terms of quizzes per course even higher than the reference client while Candidate II shows significantly lower use of these platform features.

The presence of quizzes and/or graded assignments is, however, not enough to be able to fit a risk model. These grades need to show some variance as well as some correlation to the final grades or other outcome variable. To ascertain if such a pattern exists we generate a density plot of point biserial correlation calculated between these variables of a course by course basis, for each of the clients as well as the reference data. Figure 1 shows the density of correlation between assignment grades and the outcome variable.

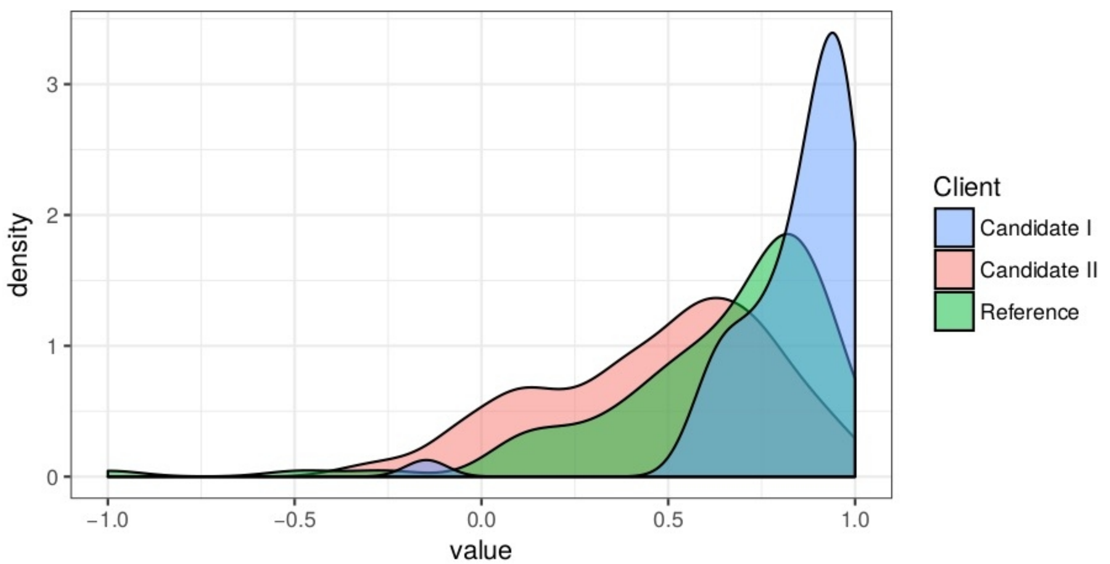


Figure 1: Density of correlation between Assignment Grade and Final Grade

We see that Candidate I has an even higher density of strong correlation between the variables than the reference institution. Candidate II shows a lower correlation overall, and, interestingly, a non-trivial portion of the density curve is found below the zero midpoint, i.e. indicates some systematic portion of negative correlation between the variables. These cases, where systematic negative correlations are found, are assumed to be invalid and are discarded for modeling purposes.

Figure 2: Density of correlation between Quiz Grade and Final Grade

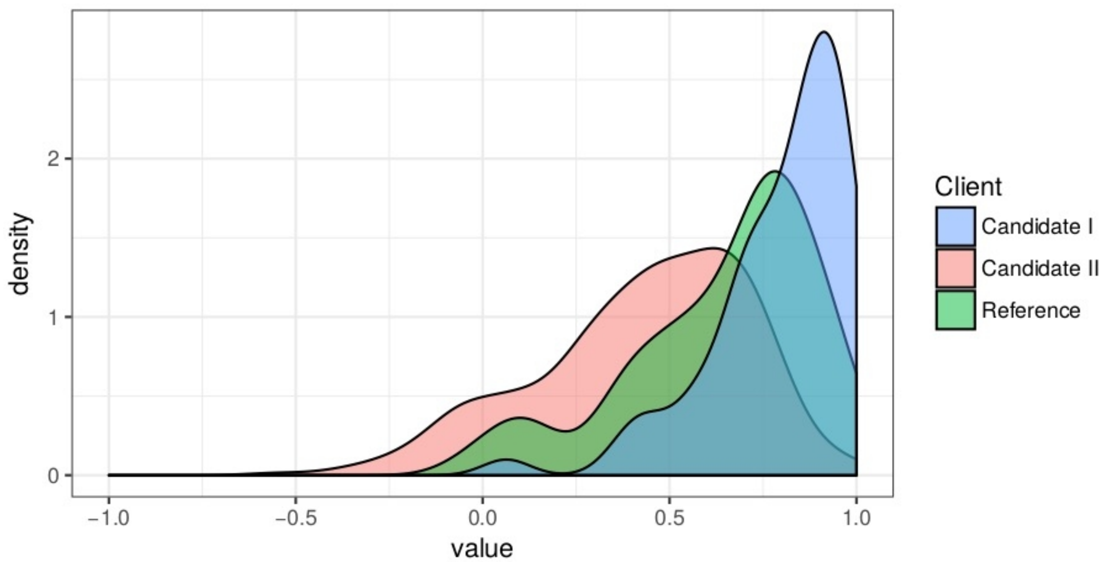


Figure 2 shows the same density plot for the correlations between quiz-grades and final grades. A pattern similar to that in Figure 1 can be appreciated, albeit with slightly lower incidence of negative correlation.

We see that Candidate I fares well. Candidate II, however, shows higher density of lower correlations (and even a substantial density of negative correlations) between the two variables, meaning that overall activity-level is not a stable predictor of success. This pattern is typically found when the institution has a higher proportion Supplemental and/or Complementary, as per the archetypes discussed previously. Based on these observations we draw the conclusion that fitting a risk model for Candidate I is likely to be successful, while Candidate II does not have enough meaningful use of the LMS for this to be the case.

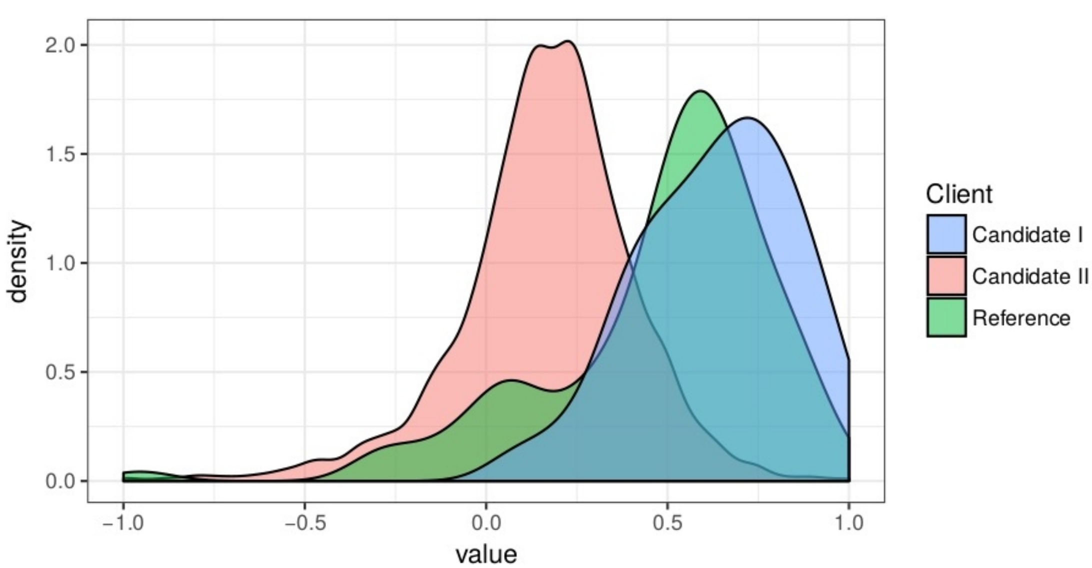


Figure 3: Density of Correlation Between Number of Clicks per Student and Final Grade

Based on these observations we draw the conclusion that fitting a risk model for Candidate I is likely to be successful, while Candidate II does not have enough meaningful use of the LMS for this to be the case.

3.3 AUTOMATIZATION

The example explored in the previous section shows that it is possible to predict the usefulness of a predictive retention risk model starting from the parameters and variables we chose. The decision to proceed or not with modeling is still, however, left up to the researchers, i.e. the very last step is still a manual one. For this procedure to become a scalable solution we need to be able to automate all the steps in the process. The application of filters as per Table 1 is trivial, but the determination of conformity of the density distribution to a reference is a bit more involved. Analysis of the data in R (R Core Team, 2016) with the fitdistrplus package (Delignette-Muller & Dutang, 2015) found that the density distribution can be modeled as a beta distribution (with α=1.732 and β=0.952). Having a theoretical distribution to test against allows us to use the Kolmogorov-Smirnov (Kolmogorov, 1933; and Smirnov, 1939) statistic as a significance test, where the null-hypothesis in that the probability density of the correlations is not significantly different from the theoretical distribution. For practical purposes it does not matter, indeed it is beneficial, if these the densities are concentrated close to the 1.0 mare, so a one-way test is appropriate. One way to visualize this is by plotting the cumulative density of each distribution alongside the theoretical one. An example of this is shown in Figure 4, where we see that the totality of the of the cumulative densities for each candidate are found on either side of the theoretical reference.

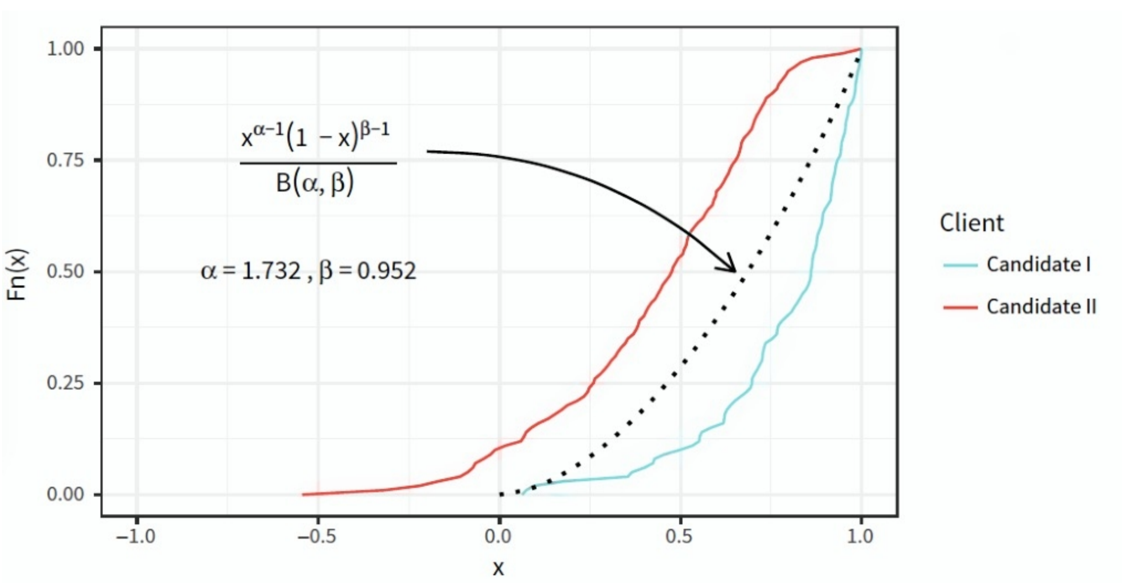


Figure 4: Cumulative Density Distribution – Theoretical and Observed

The results from the one-way Kolmogorov-Smirnov test for the two candidate institutions are shown in Table 4, and are congruent with the researchers' intuition. These results show that the procedure can be set up as a completely automated system.

Table 4: Kolmogorov-Smirnov Test for Each of the Candidates

Client	D	p-value
Candidate I	0.018	0.940
Candidate II	0.323	< 0.001

4 CONCLUSIONS AND OUTCOMES

This study shows that is it possible to quantify and predict the likelihood of a successful risk-modeling exercise based on historical data. By applying both heuristic filters and empirically extracted parameters we can avoid deploying under-performing retention risk models as well as target deployments where likelihood of success is higher.

As a result of this research, processes were put in place to pre-screen clients for risk-modeling. The X-Ray Learning Analytics product is now offered without risk modeling by default, and risk modeling is only offered where the pre-screening shows that a deployment is likely to be successful. We thus drastically reduce or even eliminate the deployment of under-performing models. At the same time we are now to identify clients for whom a deployment might be appropriate even if they are not currently using X-Ray.

5 LIMITATIONS AND NEXT STEPS

The initial filters both for institutions (population parameters) as well as course-level filters were applied based on the researchers intuition. This constitutes a limitation of the study since these precepts can and ideally should be empirically tested. The same is true for the cases where negative correlation was found between potential predictors and outcome variables. At present these are unceremoniously discarded as invalid, but it is clear that further inquiry into these marginal cases is merited as it may result in a more complete understanding of the patterns that govern and predict success in the modeling of retention risk.

REFERENCES

Blackboard Inc. (2015). X-ray learning analytics.

Campbell, J. P. (2007). Utilizing student data within the course management system to determine under-graduate student academic success: An exploratory study. In E. Studies (Ed.), Educational studies. Indiana: A. G. Rud.

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. Journal of Statistical Software, 64 (4), 1–34. Retrieved from http://www.jstatsoft.org/v64/i04/

Dietrichson, A. (2016). *Backtest of UCEM risk models* (Internal Report). Blackboard Inc.

Forteza, D. (2016). *Backtest of OSHO risk models*. Blackboard Inc.

Forteza, D., & Nuñez, N. (2016). Patterns in course design: How instructors actually use the lms. Retrieved September 20, 2017, from http://blog.blackboard.com/patterns-in-course-design-how-instructors-actually-use-the-lms/

Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. The Internet and Higher Education, 14 (2), 89–97.

Glass, G. V., & Hopkins, K. D. (1995). Statistical methods in education and psychology. Allan & Bacon.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. Giornale Dell'Institute Italiano Degli Attuari, 4 (6.1), 83–91.

Lauria, J., E. J. M. B. (2015). Mining sakai to measure student performance: Opportunities and challenges in academic analytics. Academic Conferences; Publishing International Limited.

Lopez-Raton, M., Rodriguez-Alvarez, M., Cadarso-Suarez, C., & Gude-Sampedro, F. (2014). OptimalCut-points: An r package for selecting optimal cutpoints in diagnostic tests. Journal of Statistical Software, 61 (8),1–36.

Macfadyen, L. P., & Dawson, S. (2010). "Mining LMS data to develop an 'early warning system' for educators: A proof of concept". Computers & Education, 54 (2), 588–599.

McWilliam, E., Dawson, S., & Pei-Ling Tan, J. (2008). Teaching smarter: How mining ict data can inform and improve learning and teaching practice. Graduate School of Medicine – Papers.

Morris, L. V., Finnegan, C., & Wu, S.-S. (2005). Tracking student behavior, persistence, and achievement in online courses. The Internet and Higher Education, 8 (3), 221–231.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rafaeli, S., & Ravid, G. (1997). OnLine, web based learning environment for an information systems course: Access logs, linearity and performance. ISECON.

Ryabov, I. (2012). The effect of time online on grades in online sociology courses. MERLOT Journal of Online Learning and Teaching, 8 (1), 13–23.

Smirnov, N. (1939). Estimate of deviation between empirical distribution functions in two independent samples. Bulletin of Moscow University, 2, 3–16.

Whitmer, J., Fernandes, K., & Allen, W. (2012). Analytics in progress: Technology use, student characteristics, and student achievement. EDUCAUSE Review.

Implementation of a Student Learning Analytics Fellows Program

Author(s): George Rehrey, Dennis Groth, Stefano Fiorini, Carol Hostetter and Linda Shepard
Indiana University
grehrey@indiana.edu

ABSTRACT: Post-secondary institutions are rapidly adopting Learning Analytics as a means for enhancing student success using a variety of implementation strategies, such as, small-scale, large-scale, vended products. In this paper, we discuss the creation and evolution of our novel Student Learning Analytics Fellows (SLAF) program comprised of faculty and staff who conduct scholarly research about teaching, learning and student success. This approach directly addresses known barriers to successful implementation, largely dealing with culture management and sustainability. Specifically, we set the conditions for catalyzed institutional change by engaging faculty in evidence-based inquiry, situated with like-minded scholars and embedded within a broader community of external partners who also support this work. This approach bridges the gap between bottom-up support for faculty concerns about student learning in courses and top-down administrative initiatives of the campus, such as the strategic plan. We describe the foundations of this implementation strategy, describe the SLAF program, summarize the areas of inquiry of our participating Fellows, present initial findings from self-reports from the Fellow community, consider future directions including plans for evaluating the LA research and the broader impacts of this implementation strategy.

Keywords: Institutional Learning Analytics; change management; faculty engagement; communities of transformation; student success; learning analytics fellows program

1 INTRODUCTION

Post-secondary institutions are rapidly adopting Learning Analytics (LA) to enhance student retention and graduation rates (Treaster, 2017), using a variety of approaches that show differing levels of success. Institutions are implementing both in-house early warning systems (Lonn et al., 2012; Arnold & Pistilli, 2012) as well as large-scale vended applications (such as Loud Sight, Educational Advisory Board or Civitas) and yet share common barriers to successful implementation: cultural barriers, institutional commitment, policy development, and resistance to change (Bichsel, 2012; Ferguson, et al., 2014; Macfadyen, et al., 2014). No approach appears exempt from the challenges of adoption, inasmuch as 'Institutional implementation of learning analytics calls for thoughtful management of cultural change (Macfayden, et al., 2017)."

1.1 Foundation/Rationale

In this paper, we discuss the creation and evolution of our Student Learning Analytics Fellows (SLAF) program comprised of faculty and staff who are using LA to conduct scholarly research about teaching, learning, and student success. Our main premise is that a change in faculty understanding of their students through engaged research and participation in LA development can lead to a change in institutional culture about student success. We also believe that joining a networked community of like-minded scholars provides a unique opportunity to catalyze institutional change at the course, curricular, program and institutional levels. The introduction of the SLAF program is supported by our new Center for Learning Analytics and Student Success (CLASS) and builds upon a larger community of faculty-driven work at our Center for Innovative Teaching and Learning (CITL). This includes two successful well-established programs, our FLCs (Faculty Learning Communities) and SoTL (Scholarship of Teaching and Learning) programs. Both programs acknowledge that faculty engagement is essential to the adoption of new practices, and that when successful, can lead to a change of the teaching culture at the departmental level (Austin, 2011). In the past few years, we have also come to understand the positive effects of having faculty collaborate within a larger Community of Transformation (Kezar & Gehrke, 2015), proving as an effective method to promote implementation of new models and most importantly, sustain cultural change (Fairweather, 2008; Henderson & Finkelstein, 2011). Thus, external partnerships are an integral part of our approach as we work with Bay View Alliance and other partner institutions to invest not only in LA tools, but in the people and communities that will use them as well (Bischel, 2012).

As recommended by Bischel (2012), our SLAF program aligns well with the strategic plan and objectives of the campus. Those objectives include (but are not limited to): supporting retention and graduation of students, developing best practices for recruiting and retaining diverse students, designing evidence-based curriculum, and engaging faculty in learning analytics research. This approach bridges a gap between bottom-up support for faculty concerns about student learning in courses and in the curriculum and the top-down administrative initiatives outlined in the university strategic plan, facilitating a sustained institutional change at the course, program and institutional levels. Change that 1) embraces evidence-based decision-making, 2) demonstrates an increase in faculty participation in inquiry and development of resources to support the use of learning analytics, 3) establishes sustainable faculty-led oversight including implementation of recommended activities, and 4) instills ownership for student success through the curriculum.

We now describe the Student Learning Analytics Fellows (SLAF) program at our institution, provide a summary of the faculty research to date, and the internal and external supports for this work. We also describe our collaborations with other institutions who are also adopting this approach (see LAK 2017 workshop, Macfayden, et al., 2017) and the strengths of these emerging communities. We will also summarize some of the initial evidence gathered to evaluate the success of the SLAF program and the role of these efforts within the emerging field of LA (McCoy & Shih, 2016). Given the goals of the program around institutional change, we describe future plans for the evaluation and describe the challenges that remain. With this reflection of our SLAF program, we hope to understand the broader impact of this work across our campus and enable opportunities for continuous improvement.

2 SLAF PROGRAM

The SLAF program engages faculty in the scholarship of student success. An annual call for proposals (CFP) and a campus event to explain the goals sets the stage for this program. Faculty Fellows, submit a proposal outlining their projects goals and intended outcomes and fellows with accepted proposals attend a kick-off event prior to meeting with professional Institutional Research (IR) staff to discuss their projects and to develop a research strategy. All aspects of the work are discussed with the IR staff, including the availability of data, how data will be analyzed and the skill sets of the researcher.

For some, this initial conversation is the beginning of a close partnership with the IR office while other Fellows opt to work independently, only returning to IR with specific questions or data needs.

The data required for Fellows includes individual student data about academic progress (e.g., degrees, majors, courses, grades), academic preparation (e.g., high school GPA, SAT/ACT scores), student life (e.g., residential programs, student activities), student financial status, and student demographic information (e.g., gender, ethnicity, residency). In general, longitudinal data sets and data dictionaries were provisioned that are highly structured and purposeful.

The process for provisioning data has been considered a significant obstacle that higher education faces in this emerging field (Dede et al., 2016). For this work, our approach was two-pronged, 1) linking the research proposals to the institutional mission including having administrative support and 2) administratively establishing coordination among the relevant campus compliance offices (IRB and Data Stewards). All provisioning fell within the standard protocols of each relevant office. At this scale, the process was manageable; however, with more Fellows, this may need to be revisited. For a fuller review of the provisioning processes and the implementation of the program, see Rehrey et al. (2018).

2.1 Summary of SLAF Projects

Now in its third year, 28 faculty have participated in 29 research projects, with 10 of those faculty deciding to return for a second or third year to continue their research. During the first two years alone, 24 participants, representing 11 programs, embarked upon 19 different projects. Collectively, they investigated 3.2 million student enrollment records corresponding to the career progression and characteristics of 150,000 students. The charts below provide an initial analysis of the program to date, describing the distribution of Fellows as categorized by their academic fields (Figure 1) and the student factors that the Fellows investigated as part of their research (Figure 2).

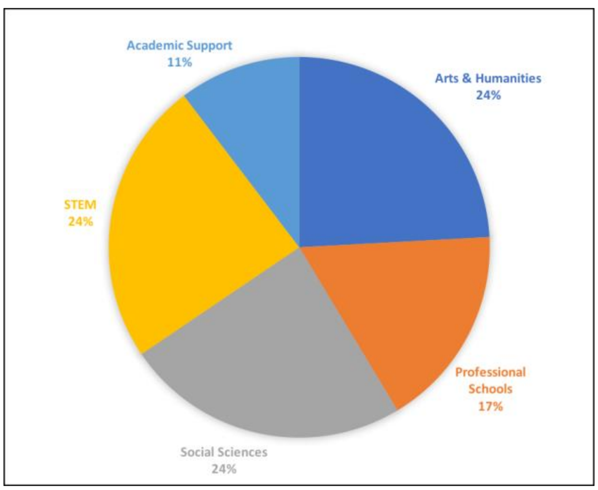


Figure 1: Academic Fields as a Percentage of 29 Fellow's Projects

A description of whether the research addresses student success at the course, curriculum/program, or university level is included as well, keeping in mind that a single project may address more than one level at a time (Figure 2).

In general, faculty projects involved inquiry within four broad categories of factors that influence student success. In many cases, the individual research projects are studying the effect of multiple factors, and at multiple levels (course, program and university). For the purpose of this initial analysis of the Fellows program, the factors were categorized in the following way: 1) *Student Demographics*: including student characteristics such as ethnicity, race, and class standing, 2) *Student Preparation*: such as transfer credits, prerequisites, curriculum pathways, pre-college courses, and remedial educational programs, 3) *Student Performance*: as understood by GPA, persistence, retention, engagement indicators and graduation rate, 4) *Student Choice*: as understood by major selection, inflection points, and pathways toward graduation.

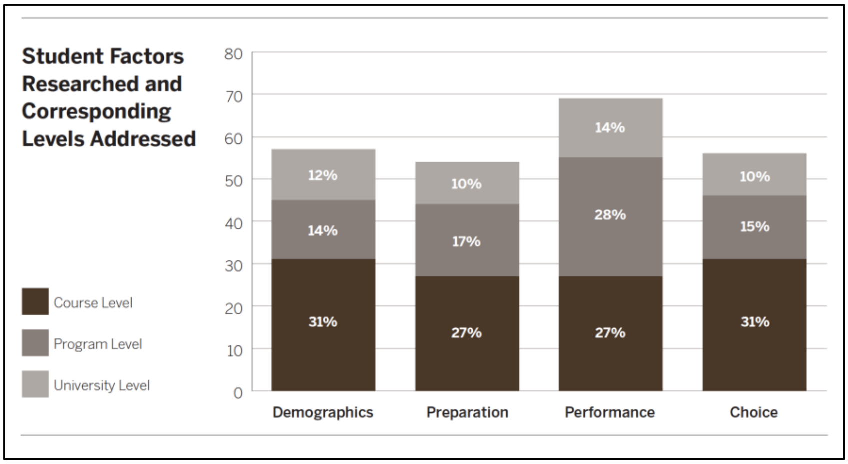


Figure 2: Fellows' research questions at course, program and institutional levels

The student factors that the Fellows are researching is fairly evenly distributed. Not surprisingly at this point in the program, faculty research tends to concentrate on the course and program levels. Over time and as research questions become more complex, this should influence the levels being addressed, with more projects making connections to data-driven institutional decision making, along with increased rates of persistence, retention and graduation at our institution.

2.2 Survey of SLAF Community – How did the program work?

The program is intentionally designed to: 1) encourage faculty to generate research questions that make actionable and pragmatic use of LA, 2) provide large, robust data sets for SoTL-type research, 3) increase the general understanding and use of LA data, 4) encourage data driven decision-making at the course, curriculum, and program levels and, most importantly, 5) change faculty perceptions about who is responsible for student success. To understand the effectiveness of the program, Likert-scale questions with open-ended comments for each question were designed for three audiences: Faculty Fellows, Sponsors of the Faculty Fellows (those who wrote a letter of support submitted with their initial proposal) and their Department/School Head. All fifteen of the Faculty Fellows responded (100% response rate), four out of nine Sponsors responded (given the low number they are not reported here) and twelve out of eighteen (67% response rate) school/department heads responded. The survey questions were approved by the university's Institutional Review Board. Surveys were sent by email with a link to Qualtrics, with a total of three requests for completion. The first research question addressed is:

> "Has LA data usage influenced programs and initiatives, and if so, how?"

The majority of the Fellows (73%) responded that working with LA data has increased the chance their departments would use data to inform decisions and that there were more conversations now about student success (60%). All Fellows had shared their projects with others in their department but were not certain that their departments have made or will make administrative decisions on the basis of their own LA projects (with only 47% answering in the "somewhat to strongly agree"). The administrators agreed that having had a participant in the LA program made their department more likely to use data to inform their decisions in the future. The next research question addressed is:

> "Has participation in the SLAF encouraged you to consider student success beyond their individual course or program?"

An interesting aspect of this question is possible selection bias wherein faculty applying for the Fellows program can be assumed to have an interest in student success. Nearly all Fellows (93%) reported that before the Fellows program, they saw student success as a part of their role as a faculty member and all of the Fellows (100%) saw student success as a part of their role as a faculty member after participating in the program. We next consider if the Fellows program helped the faculty members perceive the importance beyond their classroom and 80% agreed or strongly agreed. One Fellow made the following comment:

> "My participation in the Fellows program completely transformed me in this regard and helped turned me into a bit of a zealot for student success."

Administrators responded in a similar fashion but comments suggested that they already thought of student success as important, prior to the SLAF program. The next research question addressed is:

> "Does using LA data to conduct a research project help Fellows see the value of big data as a decision-making tool for academic decisions?"

Again, selection bias is an element to consider, so we asked the before and after questions. The majority (80%) of Fellows saw the value of using LA data to make academic decisions before the program and all of them (100%) saw the value after. One Fellow explained:

> "As the data analytics becomes more robust and easier to use, I see how data can increase teaching efficiency by providing detailed and summary feedback on content, assessment, and student learning."

Finally, we were interested in the role of community:

> "Does having a sense of being a part of a university community that had a mission for student success matter?"

While the majority (60%) were aware of other departments using data to inform their decisions, we would like to improve this figure, especially given that the majority (87%) agreed that being a part of the Fellows program helped them feel a part of a community with a mission of student success, and were now more interested in engaging in aspects of the campus community that are concerned with student success. One Fellow shows enthusiasm for the community by reflecting:

> "I found the interactions with the other Fellows and the team at (IR) to be extremely invigorating and exciting. I appreciated knowing that I am doing my work in a community of like-minded individuals in a variety of disciplines."

Conversely, another Fellow who chose to work independently remarked:

> "I believe that my process and my outcomes would have been greatly improved if I had worked with a team on this project. The ability to discuss and validate models and techniques with colleagues who know and understand this course would have been a great advantage."

Given that the SLAF program is relatively new on our campus, our knowledge of the impact is limited:

> "What administrative changes have been made as a result of the program?"

One Fellow did write:

> "Our discoveries about the importance of student motivation lead us to implement a few changes in our class curriculum."

In the actual SLAF project reports, Fellows identified many suggestions as to what their departments' administrations could do to improve student success, more follow-up on these suggestions is needed, now that more time has passed. For example, one project found that 4th year students who take a low-level course do not perform as well as the lower-level students in the same course; this suggests that the department could promote the course to lower-level students and provide tutoring services to upper-level students. Findings from another project suggest that students on academic probation who take a study skills course improve chances for retention and graduation; however, there are not enough sections for all students to take it, so a requirement is not enforced; this finding suggests allocating resources for additional sections. Other comments propose changes may take place in the future. One Fellow says:

> "Frankly, I didn't know how to do any of this before the LA grant. We had questions, but did not know how to get the answers. After the LA work, we still have a ton of questions, but we have some confidence and relationships with those who can help."

Another Fellow stated that colleagues are having more informed conversations around student success.

2.3 SLAF Community of Transformation

The broader impacts of this work are unfolding in purposeful and unanticipated ways. The connections to external communities are essential components of our implementation strategy. Four institutions within the Bay View Alliance (BVA) have become partners for this work and are implementing similar programs at their institutions (Macfayden, et al., 2017). Locally, each campus partner forms a Community of Practice whereas these broader communities (Communities of Transformation) are expected to contribute to sustained change on our campus (Kezar & Gehrke, 2015). Since these partnerships are newly formed, the outcomes of these communities are still unfolding. Another recent development on our campus is the formation of the Center for Learning Analytics and Student Success (CLASS), charged with furthering the campus' commitment to the scholarship of teaching and learning in conjunction with cutting-edge research in learning analytics to support student success. CLASS will bring together communities, including our external communities, engaged in LA work to consider knowledge gained, innovations and adoption of strategies. An unanticipated outcome of the work is an organically formed community of Fellows who are requesting funding for the formation of an Educational Data Science program, a new interdisciplinary field of study that would advance this type of work on our campus and share knowledge more broadly with relevant disciplinary communities.

3 CONCLUSION

The complexity of evaluating the impact of communities of practice (within our institution) as well as communities of transformation (extending to partnerships) is a recognized challenge. Communities are complex systems, they take considerable time (seven years) to mature (Kezar & Gehrke, 2015) and concepts around sustainability and transformation (for example) are ill-defined. Despite these challenges, we will continue to develop strategies for the evaluation of our program and focus attention on continuous improvement. Our institutional data, retention and graduation, can provide initial benchmarks of our success. Knowledge gained from self-reports from our Fellows community will continue to provide valuable feedback and we anticipate that reflections from our broader community of partners will provide valuable information as well. In addition, we have gained insights from researchers (McCoy & Shih, 2016) evaluating our work directly, as a case study. As part of this presentation, we look forward to exploring the topic of evaluation with the community at LAK 2018.

As we consider the rich faculty communities (FLC and SoTL) that provided the foundation for the SLAF program we recognize that these influences move in two directions. SLAF programs will continue to be influenced by these communities and these communities will be influenced by LA and the SLAF community. We anticipate growth for all, with greater capacity to improve teaching, learning, and student success at our institution and within higher education. This was recognized when the SLAF community on our campus came together to present final projects and reflect on our first year. Our esteemed SOTL scholar, Craig Nelson states (at the Student Learning Analytics Showcase, November 19, 2015):

> "It's been clear for a long time that every class is an experiment and one for which we traditionally throw the data away as soon as we gather it. And what's clear here is there's also an immense amount of external data that has not been routinely brought to bear in any thinking. And that the learning analytics is going to make it easier to look at the external data and in the process motivate people to look a lot more systematically at the internal data. So, I am very cheered by all of this."

The boundaries of LA are expansive and the potential for LA to enhance student success remains

> "the most dramatic factor shaping the future of higher education (Siemens & Long, 2011)."

REFERENCES

Pistilli, M. D. and K. E. Arnold (2012). *Course signals at Purdue: Using learning analytics to increase student success.* In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Number May, New York, NY, USA: ACM. LAK' 12., pp. 2–5.

Austin, A. E. (2011). *Promoting evidence-based change in undergraduate science education.*: Fourth Committee Meeting on Status, Contributions, and Future Directions of Discipline-Based Education Research.

Bischel, J. (2012). *Analytics in Higher Education: Benefits, Barriers, Progress and Recommendations*. Retrieved from Louisville: http://www.educause.edu/ecar

Dalrymple, M. S., L. (2016). *Engaging Faculty in the Scholarship of Student Success*. Paper presented at the Association for Institutional Research Forum, New Orleans, LA.

Dede, C. H., Andrew; Mitros, Piotr. (2016). Big Data Analysis in Higher Education: Promises and Pitfalls. *Educause*(September-October 2016), 23-34.

Fairweather, J. (2008). *Linking Evidence and Promising Practices in Science, Technology, Engineering, and Mathematics (STEM) Undergraduate Education* Retrieved from Washington, DC:

Felten, P. J. (2013). Principles of Good Practice in SoTL. *Teaching and Learning Inquiry: The ISSOTL Journal, 1*(1), 121-125.

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). *Setting learning analytics in context: overcoming the barriers to large-scale adoption*. Paper presented at the Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, Indianapolis, Indiana, USA.

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching, 48*(8), 952-984. doi:10.1002/tea.20439

Hutchings, P., Huber, M. T., & Ciccone, A. (2011). *The Scholarship of Teaching and Learning Reconsidered: Institutional Integration and Impact*: Wiley.

Kezar, A., & Gehrke, S. (2015). *Communities of transformation and their work scaling STEM reform.*: Pullias Center for Higher Education, Ross School of Education, University of Southern California.

Lonn, S. et al. (2012) *Bridging the gap from knowledge to action: Putting analytics in the hands of academic advisors*, LAK12:2nd International Conference on Learning Analytics and Knowledge*,* 3-4, Vancouver, BC.

Macfayden, L. P. D., Shane; Pardo, Abelardo; Gasevic, Dragan. (2014). Embracing Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Challenge. *Research & practice in assessment, 9*(2), 17-28.

Macfayden, L. P. S., M.; Rehrey, G.; Shepard, L.; Greer, J.; Groth, D.; Molinaro, M. (2017). *Developing institutional learning analytics "communities of transformation" to support student success*. Paper presented at the Learning Analytics and Knowledge Conference, Vancouver, BC. https://doi.org/10.1145/3027385.3029426

McCoy, C., Shih, Patrick. (2016). Teachers as Producers of Data Analytics: A Case Study of a Teacher-Focused Educational Data Science Program. *Journal of Learning Analytics, 3*(3), 193-214.

Rehrey, G. G., D.; Hostetter, C.; Shepard, L. (2018). The scholarship of teaching, learning, and student success: Big data and the landscape of new opportunities. *Conducting & applying the scholarship of teaching and learning beyond the individual classroom level.* Bloomington, IN: Indiana University Press.

Rehrey, G. S., Greg; and Hostetter, Carol. (2014). SoTL Principles and Program Collaboration in the Age of Integration. *International Journal for the Scholarship of Teaching and Learning, 8*(1), 1-11. doi:https://doi.org/10.20429/ijsotl.2014.080102

Siemens, G. & Long, P. (2011). Penetrating the analytics in learning and education. *Educause review, 46*(5) 31-40.

Treaster, J. B. (2017, 2/2/2017). Will you graduate? Ask big data. *The New York Times*. Retrieved from https://www.nytimes.com/2017/02/02/education/edlife/will-you-graduate-ask-big-data.html?_r=1

Analysing student responses: early lessons from a pilot study

Jenny McDonald
University of Auckland
j.mcdonald@auckland.ac.nz

Adon Christian Michael Moskal
Otago Polytechnic
adon.moskal@op.ac.nz

Irina Elgort
Victoria University of Wellington
irina.elgort@vuw.ac.nz

Cathy Gunn
University of Auckland
ca.gunn@auckland.ac.nz

ABSTRACT: We report on early lessons from a pilot study to evaluate a new web-based tool for teachers. The tool is designed to support the rapid analysis of written student responses to short answer questions and was conceived to support formative assessment, especially in large-class settings, as well as to provide insights into teaching and learning design. We describe our approach to building the tool through working in partnership with academic developers and teaching staff from diverse academic contexts. We then discuss the challenges and opportunities that this process has presented. Insights from the pilot study to date suggest that the affordances of existing NLP technologies can be deployed to the advantage of teachers, and ultimately learners, by making the analysis of student responses in a range of contexts easier, quicker, more robust and transparent.

Keywords: Formative assessment, text analytics, automated short answer question marking, educational technology development.

1 INTRODUCTION

Beyond summative assessment of written student responses to questions, the rapid analysis of student responses has potential not only to provide opportunities for formative assessment but also to help teachers to assess when and in what ways teaching and student learning can be enhanced (McDonald, Bird, Zouaq & Moskal, 2017). Text is arguably central to teaching and learning; from deep questions designed to reveal how students understand and describe the world, to more surface-level questions relating to content knowledge. In addition, students' communicative ability is often assessed through examination of syntax and style. Yet, particularly in large class settings, it is challenging for educators to know what and how their students are learning. The time and resource constraints of modern tertiary environments mean it is often only at the time of marking final examinations that student understanding becomes clear to the teacher, by which time it is too late to respond to misconceptions.

1.1 Background

Automated assessment of student responses is one approach to addressing these constraints; to achieve outputs which are on a par with human graders is both an active research area and a work in progress (Burrows, Gurevych, & Stein, 2015). However, the goal of fully automated assessment can shift depending on how the assessment is conceived and structured. For example, assessing short responses to questions designed to check recall of facts is an easier task for both humans and machines than assessing longer responses to deeper or open-ended questions (Dzikovska, Nielsen & Leacock, 2015). Formative assessment involves more than assigning a single label or grade to a student response. In practice, multiple labels may be required to adequately categorise a single response (McDonald, Bird, Zouaq & Moskal, 2017). Reliability, both between human graders and within a single grader is also an issue (Elton & Johnston, 2002; Jonsson & Svingby, 2007)—if humans struggle to ascribe a consistent meaning to a response, even with the aid of scoring guides and rubrics, it is hard to see how an automated system will fare better. Further, in higher education settings, teachers who take an interpretivist approach may contest the notion of reliability itself (Orr, 2007). Finally, good teaching practice dictates that assessments can and should change regularly which adds a further layer of complexity for automated assessment, where supervised methods are used, since training data will be limited.

To address some of these issues, recent studies have adopted a more nuanced approach. Automatic methods are used to support human decision-making rather than replace it (e.g. Basu, Jacobs & Vanderwende, 2013). This provides several advantages. There is potential to improve marker consistency and objectivity as well as portability between contexts (Pado & Kiefer, 2015). Teaching insight can be gained and pedagogenic errors—errors induced by the teacher or by the teaching environment (Laurillard, 2002)—identified by looking for co-occurrence of text fragments between responses and between responses and teaching materials (McDonald, Bird, Zouaq & Moskal, 2017). A key advantage is that the analytic process, and indeed the learning process, represented by student text is made more transparent for the teacher. This not only helps to assuage concerns about 'black-box' algorithms replacing human input (in particular where high-stakes assessment is concerned, e.g. Ericsson & Haswell, 2006; Pado & Kiefer, 2015), it also provides a window into student understanding, and thus the opportunity for a *teaching moment* (Havighurst, 1952).

1.2 A text analysis tool for teachers

Consistent with the use of text analytic methods to support human assessment we describe a pilot study to evaluate a new text analysis tool for teachers. Quantext is an online platform designed to help teachers to extract insights from student responses to short-answer questions (McDonald & Moskal, 2017). While the following description outlines the key features of Quantext in terms familiar to natural language processing (NLP) researchers and practitioners, it is important to note that the Quantext interface has been designed for the non-specialist. The use of jargon is avoided, and where it is used, is explained through in-context tooltip displays.

Teachers upload student responses to Quantext, which extracts text features from the dataset, and aggregates and presents them in a variety of forms for further exploration. For example, Quantext first displays the most frequent words, bigrams (two word units) and trigrams (three word units) from the student responses. Teachers can then click on a word or multi-word unit of interest, and explore how it is being used by students via a keyword-in-context display and associated wordtree visualisation (Wattenberg & Viégas, 2008). Quantext allows teachers to label responses via a sorting process based on selected features such as response length, words or ngrams (multi-word units), as well as on readability indices and semantic similarity. Labels are created by teachers and multiple labels can be applied to any given response. Text-based teaching materials, such as lecture transcripts, handouts or course books can be uploaded as a reference corpus and features common to both student and teacher discourse can be highlighted. Stopwords, choice of common readability indices, algorithms for calculating ngram keyness, and semantic similarity measures are all in control of the teacher. Finally, Quantext has been designed specifically to support comparison of analyses both within and between student cohorts.

In the next section, we describe our approach to building the tool through working in partnership with academic developers and teaching staff from diverse academic contexts. Specific design decisions arising from this partnership are highlighted.

2 APPROACH TO QUANTEXT DEVELOPMENT – A PILOT STUDY

2.1 Starting from student responses

The need for a tool like Quantext became apparent through a series of New Zealand-wide workshops designed to introduce existing text analysis tools to teachers. The workshops were part of a NZ-wide Ako Aotearoa funded project, *Building an evidence-base for teaching and learning design using learning analytics* (Gunn & McDonald, Forthcoming). Teachers and learning designers who attended the workshops expressed enthusiasm for analysing text and a willingness to try existing and readily available tools (e.g. http://www.laurenceanthony.net/software/antconc/ (Anthony, 2014), http://www.sketchengine.co.uk (Kilgariff et al., 2014), and Textalyser.net). However, while in principle the idea was well-received, in practice it was clear that existing text analytic tools presented numerous obstacles for teachers unfamiliar with concepts spanning linguistics, computing, data management and statistics. From feedback during the workshops, it was clear that the design and purpose of existing tools was not well-aligned with the specific needs of most teachers. This is not a criticism of the tools themselves, far from it; rather, it is merely a reflection of the fact that they were not designed to meet the specific needs of teachers seeking insights from student generated text. We therefore directed our efforts towards learning from the experience and creating a tool geared towards the average teaching academic. Informed by the findings from an exploratory study situated in an undergraduate health sciences context (McDonald, Bird, Zouaq & Moskal, 2017), we created a proof of concept in the form of a Jupyter notebook. While still far from accessible to most teachers, the basic functions of the notebook involved iterative sorting of student responses based on simple text features such as response length and key words and visualising the results. This approach resonated with academic developer colleagues and from there the first iteration of Quantext as a web-based tool developed.

2.1 Working in partnership with teaching staff

While still in a rudimentary stage, and in partnership with academic developers who work within tertiary institutions to support teacher professional development, we recruited a small number of teachers to a pilot study of the fledgling tool. The 8 pilot participants so far have been drawn from 4

NZ tertiary institutions; 3 universities and 1 polytechnic. Both the teachers, and their academic developer guides, provided early evaluation and input into the iterative development of the tool. Tertiary teachers involved in the pilot are from diverse disciplines with the majority interested in analysing text from their undergraduate classes. The range of disciplines includes Physics, Philosophy, Architecture and Design, and Medicine with class sizes ranging from around 100 students to more than 2,000. In addition, some pilot participants were interested in trying out Quantext with data from Massive Open Online Courses (MOOCs)—one is a statistics MOOC offered through the University of Auckland (approximately 20,000 students at the start of the course), while the other is a MOOC on Antarctica offered through Victoria University of Wellington (approx. 2,000 students at the start of the course – Elgort, Lundqvist, McDonald & Moskal, 2018). Students may join or leave the MOOCs at any time.

A key feature of the pilot is that while all teachers came to it with an interest in analysing student text, they also came with questions, assignments, student responses from earlier cohorts, and in some cases discussion forum posts; data specific to their context and not designed to test the tool. In other words, the context came first rather than tool development. In this way we planned to iteratively refine the tool in order to address specific teaching needs. At the time of writing the pilot study is still underway and planned to continue through semester 1, 2018. The broad goals of the pilot study are to evaluate the utility of Quantext along several dimensions (e.g. speed of analysis, validity and reliability of data and so on) and report on teacher reflections from tool use. While still at an early stage, we have already identified challenges and opportunities revealed by the pilot study to date. We describe these challenges and opportunities in the final sections of this paper.

3 CHALLENGES

3.1 Data input

We knew from interviews with teachers, and survey results conducted as part of the Ako Aotearoa project, that getting data into and out of systems often presents a challenge for teachers. For the purpose of the pilot study we therefore standardised the data input format to a Microsoft Excel spreadsheet—questions are listed in the first worksheet of the spreadsheet, one question per row; student ids and responses are listed in subsequent worksheets which are numbered according to their corresponding question number. While in the short-term this simple approach has resulted in some calls for help, in the main any issues have been quickly resolved and have the benefit that once uploaded, all data is in a consistent format. Eventually, input of question and response data should be directly integrated with Learning Management Systems (LMS) and other common assessment platforms.

3.2 Question length and style

An associated challenge was handling the range in style of question and lengths of responses. Question style so far has ranged from single questions to multiple part and sub-part questions. The ability to retain context across related questions is important and needs to be handled. In our earlier work, student responses had been less than 50 words in length with an average response length of around ten words. By contrast, the average response length from early pilot data was closer to 200 words. While this was not an issue in terms of data format for uploading, or for our processing engines, it did prove an issue for how to display responses in the fledgling Quantext interface. In particular, our spreadsheet view became unwieldy, as these longer responses resulted in excessive scrolling down the page to view all responses. These issues are currently being addressed.

3.3 Specialist terminology

Although we took care to reduce specialist terminology wherever possible in the Quantext interface, even the most basic terms in everyday use can become problematic when they also have a specific technical meaning. For example, in corpus linguistics, a 'keyword' is one which occurs more often than expected by chance and is calculated by comparing word frequency between a given text and a reference corpus. In common use, 'keyword' simply means a word which represents the central meaning of a text. It became apparent from discussions early in the pilot that these differences need to be made explicit, or alternative terms chosen, in order to avoid confusion. Key phrases and blacklist words are other examples.

3.4 Accessible student response corpora

We found one of the most useful things, when recruiting participants to the pilot study, is to demonstrate the tool with an authentic dataset and one which resonates with participants. Furthermore, such datasets, in particular ones which include teacher annotations or categorization, are invaluable to evaluate both interface usability and unsupervised metrics such as similarity scoring. However, there are few publicly accessible datasets available for this purpose. One example is of student responses to questions in an undergraduate computer science course (Mohler & Mihalcea, 2009), and a second contains questions relating to understanding of US civics (Basu, Jacobs & Vanderwende, 2013). There are also some limited datasets available on Kaggle (https://www.kaggle.com/datasets). In practice, we have found far greater variability in terms of question format, number of responses and response length from the data we have observed in the pilot study thus far than from existing publicly available datasets. We are therefore seeking permission to release (anonymised) datasets obtained through this project to the wider research community.

3.5 Cost of development and support

Development and support costs can present a challenge. To some extent we have overcome these issues through developing Quantext as a sideline to our day-jobs. We hope that by releasing Quantext as an open source project this will attract wider interest and additional resource. In the meantime, alignment of pilot study goals with individual teacher/academic developer research interests has contributed to keeping costs down, however additional resource will be required to substantially progress the project.

Time and workload are further costs. Tertiary teachers have many demands made of them daily. Contributing to any pilot study takes time and of necessity is fitted in around other work. Learning to use a new tool and contribute feedback on it may result in tasks taking longer than they otherwise would. This can be counterproductive; in the early stages, a tool designed to save time may paradoxically take more time to use. For this reason, among others, pilot studies such as this one inevitably tend to recruit and indeed rely on, highly motivated teachers who may not necessarily represent the wider user cohort. This can have an impact on wider adoption (e.g. Gunn, Woodgate & O'Grady, 2005).

3.6 Integration with institutional systems and LMS

Consistent feedback from pilot study participants has pointed to the desirability of developing plugins to use Quantext within institutional LMS or leveraging LMS APIs to automatically populate Quantext. The ideal is for teachers to set questions within an LMS or other assessment tool and have student responses automatically available in Quantext. Following analysis it makes sense to export categorised output directly to tools such as the Student Relationship Engagement System - SRES (Liu, Bartimote-Aufflick, Pardo & Bridgeman, 2017) which would facilitate automating feedback, based on assigned labels, directly to students. LMS integration will almost certainly reduce the length of time taken for analysis. A goal of the pilot is to evaluate the speed of analysis independently of LMS integration. Assuming a positive evaluation, we plan to implement integration enhancements at the conclusion of the pilot project.

4 OPPORTUNITIES

4.1 Professional teacher development

The importance of academic developers working with teachers to explore the use of Quantext (or arguably any learning analytic tool) in its early stages of development, cannot be overstated. This assists with ensuring teacher and learner needs are appropriately addressed and the potential for teaching improvement and enhancing formative assessment is realised. There are opportunities to help teachers with analysing textual data in general and with writing effective short answer questions. For example, ambiguous questions can be quickly identified when the most frequent words or multi-word units in responses turn out to be completely unexpected. Furthermore, comparison of student responses with teaching materials allows teachers to see directly the impact of some of their teaching choices and learning designs reflected in student responses. This in turn encourages reflection and action (Schön, 1987).

4.2 Analytics for online fora and student evaluations

As a result of requests for functionality arising from the pilot we need to decide whether to extend Quantext to handle other forms of text-based data common in a higher education context. For example, currently Quantext is ill-equipped to handle online discussion forum data; Quantext treats student responses as independent, but forum data is highly dependent on the relationships between posts (e.g. post order, or whether a post is a direct reply to another). Incorporating such data into Quantext, other than by treating posts as independent responses (Elgort et al., 2018), will necessitate changes to the way in which data is stored and referenced, and changes to the interface to display the relationships inherent in the data. Another common data source in teaching and learning contexts are the free text comments sections of student evaluations and surveys. There is also scope to explore the use of Quantext as a research tool for analyzing qualitative elements of surveys and interview data. While Quantext can currently process this data and present teachers with summary statistics, additional NLP techniques such as sentiment analysis may be useful additions to the featureset.

5 CONCLUSION

End-user text analytic tools and platforms must be able to deal flexibly with data sources, be robust, easy to use and fast (Ittoo, Nguyen, & van den Bosch, 2016). Quantext for academic contexts currently aligns well with these aims. In addition, the analysis of student text can and should benefit from the input of teachers. Early indications from this pilot study suggest that there is no need to wait for NLP approaches to fully automated processing to achieve comparable accuracy to human markers, if indeed this is desirable in practice and can be achieved. The affordances of current advances in NLP technologies can be deployed to the advantage of teachers and ultimately learners by making the analysis of student responses in a range of contexts easier, quicker, more robust and transparent.

REFERENCES

Anthony, L. (2014). AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/

Basu, S., Jacobs, C., & Vanderwende, L. (2013). Powergrading: A clustering approach to amplify human effort for short answer grading. *Transactions of the Association for Computational Linguistics,* 1, (pp. 391–402).

Burrows, S., Gurevych, I., & Stein, B. (2015). The eras and trends of automatic short answer grading. *International Journal of Artificial Intelligence in Education, 25*(1), (pp. 60-117).

Dzikovska, M. O., Nielsen, R. D., & Leacock, C. (2015). The joint student response analysis and recognizing textual entailment challenge: Making sense of student responses in educational applications. *Language Resources and Evaluation*, 1–27.

Elgort, I., Lundqvist, K., McDonald, J., & Moskal, A.C.M. (2018). Analysis of student discussion posts in a MOOC: Proof of concept. *Companion Proceedings 8th International Conference on Learning Analytics & Knowledge (LAK18).*

Ericsson, P. F., & Haswell, R. H. (2006). *Machine scoring of student essays: Truth and consequences*. Utah State University Press.

Elton, L & Johnston, B (2002) *Assessment in Universities, a critical review of research*. LTSN Generic Centre. Available online at https://eprints.soton.ac.uk/59244/1/59244.pdf

Gunn, C., & McDonald, J. (in press). Promoting learning analytics for tertiary teachers: A New Zealand case study, In J. Lodge, L. Corrin & J. Cooney Horvath (Eds.), *Learning analytics in the classroom: Translating research for teachers.* Routledge.

Gunn, C., Woodgate, S., & O'Grady, W. (2005). Repurposing Learning Objects: A Sustainable Alternative? *Association of Learning Technology Journal, ALT-J, 13*(3), (pp. 189-200).

Havighurst, R. J. (1953). *Human development and education*. Oxford, England: Longmans, Green.

Ittoo, A., Nguyen, L.M., & van den Bosch, A. (2016). Text analytics in industry: challenges, desiderata and trends. *Computers in Industry, 78*., (pp. 96-107).

Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational research review*, *2*(2), (pp. 130-144).

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovvář, V., Michelfeit, J., Rychlý, P & Suchomel, V. (2014) The Sketch Engine: ten years on. *Lexicography*, 1. (pp. 7-36)

Laurillard, D. (2002) *Rethinking University Teaching: a framework for the effective use of educational technology*. RoutledgeFalmer, London.

Liu, D.Y.-T., Bartimote-Aufflick, K., Pardo, A. & Bridgeman, A.J. (2017). Data-Driven Personalization of Student Learning Support in Higher Education. In Peña-Ayala, A. (Ed.) *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning*. Springer. (pp. 143–169)

McDonald, J., Bird, R.J., Zouaq, A. & Moskal, A.C.M. (2017) Short answers to deep questions: supporting teachers in large-class settings. *Journal of Computer Assisted Learning* 33(4), (pp. 306–319). doi: 10.1111/jcal.12178

McDonald, J. & Moskal A.C.M. (2017). Quantext: analysing student responses to short-answer questions. In H. Partridge, K. Davis, & J. Thomas. (Eds.), *Me, Us, IT! Proceedings ASCILITE2017: 34th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education* (pp. 133-137).

Mohler, M., & Mihalcea, R. (2009*).* Text-to-text semantic similarity for automatic short answer grading*. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 567-575). Association for Computational Linguistics.

Orr, S (2007) Assessment moderation: constructing the marks and constructing the students, *Assessment & Evaluation in Higher Education*, 32(6), (pp. 645-656), DOI: 10.1080/02602930601117068

Pado, U. & Kiefer, C. (2015). Short answer grading: When sorting helps and when it doesn't. *Proceedings of the 4th workshop on NLP for Computer Assisted Language Learning at NODALIDA 2015*. NEALT Proceedings Series 26 / Linköping Electronic Conference Proceedings 114 (pp. 42–50).

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.

Wattenberg, M. & Viégas, F.B., (2008). The word tree, an interactive visual concordance. *IEEE transactions on visualization and computer graphics*, *14*(6)

Analysis of student discussion posts in a MOOC: Proof of concept

Irina Elgort
Victoria University of Wellington
irina.elgort@vuw.ac.nz

Karsten Lundqvist
Victoria University of Wellington
Karsten.Lundqvist@vuw.ac.nz

Jenny McDonald
University of Auckland
j.mcdonald@auckland.ac.nz

Adon Christian Michael Moskal
Otago Polytechnic
adon.moskal@op.ac.nz

ABSTRACT: Students' communications in a Massive Open Online Course (MOOC) may offer unique insights into their thinking and engagement with course topics. Is it possible for MOOC instructors to access such insights, short of reading individual posts? This proof of concept study used an online text analysis tool for teachers, *Quantext*, to examine key topics and ideas expressed by students in discussion forums, in a science MOOC about Antarctica. Outputs of the basic text analysis were scrutinized, in order to identify key topics of interest in the introductory MOOC forum and trace them to in-course topic discussions. We found that the analysis was, at the very least, minimally adequate to quickly observe lexical patterns in students' writing, linking their pre-course interests and aspirations to some aspect of their in-course communications.

Keywords: text analytics, MOOC discussion forums, student-generated text

1 INTRODUCTION

As numbers of Massive Open Online Course (MOOC) offerings continue to grow, education researchers investigate this novel learning environment in order to understand factors that could improve its effectiveness. Because of the sheer scale of MOOCs, their student populations are much more diverse than in traditional online courses, making it challenging for instructors to understand learning expectations and motivations of their students and keep their finger on the learning pulse throughout the course.

These issues are being addressed by using learning analytics to identify and understand patterns in students' course behaviour, aiming to reinforce those that are associated with successful learning outcomes and counter those that lead to disengagement and dropping out. The types of data used in education research and interventions are generally either clickstream data providing a detailed account of students' interactions with different components of the course (Balakrishnan & Coetzee, 2013; Boyer & Veeramachaneni, 2015; Sharma, Jermann, & Dillenbourg, 2015) or participation and achievement data from formative and summative assessment activities (Beheshtiha, Hatala, Gašević, & Joksimović, 2016). Another relevant type of data is students' interactions with peers and instructors, usually, in discussion forums or through peer assessment (Jiang, Williams, Schenke, Warschauer, & O'Dowd, 2014; Reich, Tingley, Lede-Luis, Roberts, & Stewart, 2015; Wang, Wen, & Rose, 2016). Peer interactions are often examined using social-network analysis (Jimoyiannis & Angelaina, 2012; Piech et al., 2013) identifying the degree of social interactions and centrality of individual students in online learning communities (Joksimović et al., 2016; Poquet & Dawson, 2016). There are fewer examples of learning analytics tools that support the examination of qualitative data generated by students as part of the learning process, such as answers to open-ended questions or course discussion posts. And yet, this data might provide insights into cognitive and metacognitive processes that underpin student engagement, motivation and, ultimately, their learning (Crossley, Paquette, Dascalu, McNamara, & Baker, 2016; Kovanović et al., 2016, McNamara, Allen, Crossley, Dascalu, & Perret, 2017; Wen, Yang, & Rosé, 2014).

It is unreasonable to expect instructors to read and engage with individual posts from hundreds or even thousands of students participating in a MOOC, therefore, tools and approaches are needed to allow MOOC instructors to engage with this rich and complex text data in a meaningful, effective and efficient way. Natural language processing (NLP) has been a key approach to this challenging issue. NLP tools use computational analyses of linguistic properties of student-generated text for topic modelling, generating computational indices associated with topic comprehension, high-order thinking, engagement, emotional state and motivations. Together with student behaviour and achievement data, these indices may be used to improve the quality of learning in MOOCs (Crossley et al., 2016, Wang et al., 2016).

NLP-based learning analytic tools, however, are still relatively new and more research is needed to interpret indices they generate with confidence (McNamara et al., 2017). One approach to overcome this limitation, while taking advantage of the affordances of automated text analysis, is to reveal patterns of student language use to their MOOC instructors, leaving the final interpretation and categorisation of these patterns to them. In this paper, we explore the potential of a new text analysis tool for teachers, *Quantext* (McDonald & Moskal, 2017), featuring a simple and intuitive online user interface to serve as such a tool.

2 PRESENT STUDY

2.1 Quantext

The initial Quantext analysis is conducted in three simple steps by uploading a spreadsheet containing prompts and responses, selecting a prompt to analyse, and running the analysis. This displays basic descriptive statistics and charts for each question-based dataset, including number of responses, mean response length in words and sentences, most frequent words and multi-word units (bigrams or trigrams), and readability indices (Figure 1). By selecting a frequent word or multi-word unit, users can quickly access student original texts in the worksheet view.

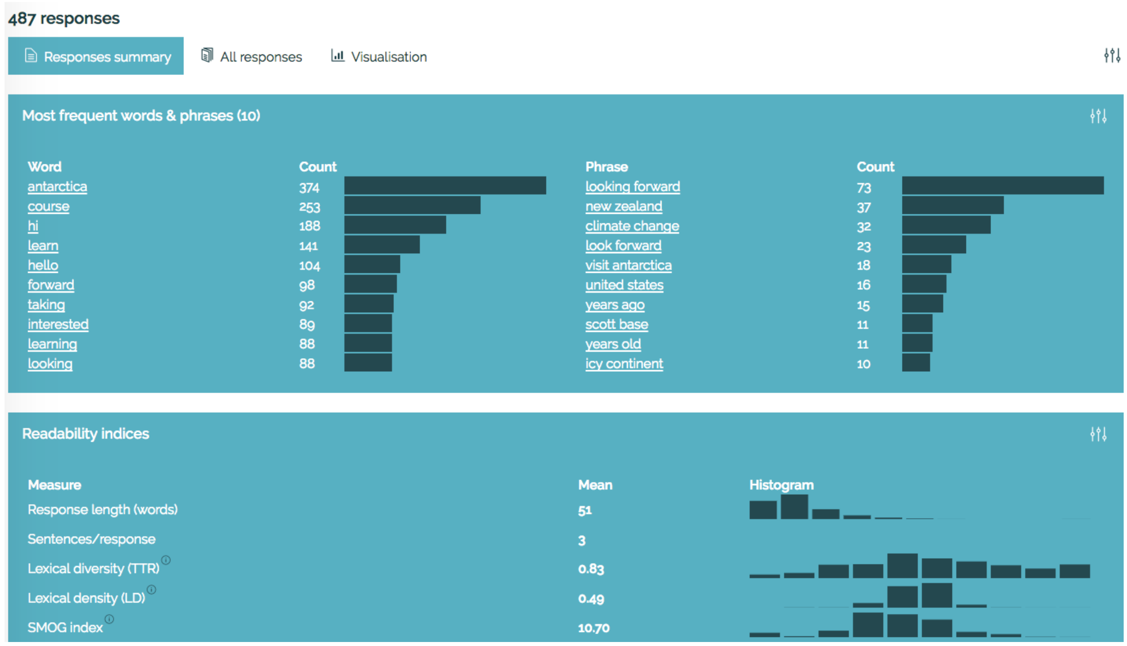


Figure 1: Quantext analysis: Responses summary

2.2 Proof of concept

We report on outcomes of a proof of concept study that evaluated whether the basic Quantext analysis is sufficient to quickly engage with student-generated textual outputs in a MOOC. Two sets of student-generated data are considered: (1) posts in a pre-course introductory forum, and (2) discussion forum posts generated in response to instructor questions linked to specific course topics. By examining students' text, framed by the Quantext analysis, we investigate whether a non-linguist would be able to identify meaningful patterns that can inform teaching.

2.3 The MOOC

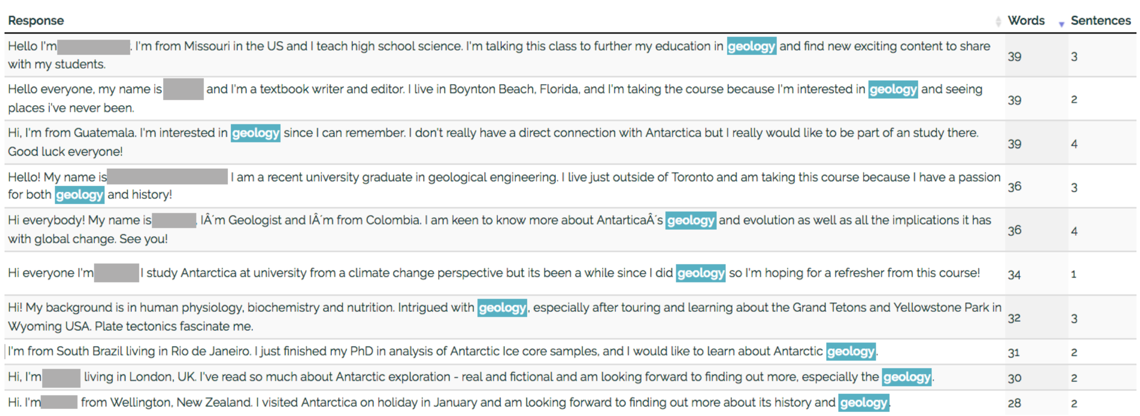
We used an inaugural Victoria University of Wellington MOOC, *ICE101x Antarctica: From Geology to Human History*, hosted on the edX platform. The MOOC was launched in April 2017 as a 5-week course. First two weeks are focused on the human history of Antarctica including exploration and the science done in Antarctica. The topic of weeks three and four is Antarctic geology, including research into Antarctica's historical climate change. The final week is about modern Antarctica, in which students meet science and humanities researchers studying the icy continent.

Course content is delivered primarily via video lectures broken into 3-8 minutes segments some of which are followed by non-assessed knowledge-check questions. At several points in the course, students are invited to respond to questions related to the topic they are studying, using discussion forums. It is explained that the forums are an opportunity to discuss and explore course topics in more depth. The forums are moderated but instructors do not engage with students in the forums. Instead, they address some of the topics raised in the forums in their weekly blog posts. There were 2161 students who signed into the course, of which 2020 went beyond the entry page. The number of active MOOC participants changed from 1757 in the first week to 510 students in week five.

3 ANALYSIS

3.1 A pre-course forum case

Our first question is what can be learned from the analysis of students' posts to the pre-course introductory forum. In this forum, students were invited to introduce themselves, share reasons for taking the course and elaborate on their connections with Antarctica. There were 487 posts in this forum from 437 students - about one fifth of those enrolled in the course (21.6%). The results of the basic Quantext analysis show a number of themes immediately identifiable by eyeballing the most frequently-used words and multi-word units. The main overall focus of the students' interest was Antarctica - the most frequently-used content word in this forum (n=374). Among the frequently-used bigrams we see: *Scott base*; *icy continent*; *South Pole*; *polar regions*; *Ross Sea*; *Antarctic peninsula* and *amazing continent*. These bigrams are aligned with the main topic of the course. *Visit Antarctica* was the fifth most frequent bigram, while *bucket list* also appeared among the 30 most frequent bigrams, pointing to a possible source of intrinsic motivation for enrolling in the course. Frequently-used trigrams were indicative of the students' attitudes to the topic and their intention to learn: *fascinated by/with Antarctica*; *interest/ed in Antarctica*; *connection with/to Antarctica*; *learn/ing about Antarctica*; *knowledge of Antarctica*. Another prominent point of interest for students enrolled in *ICE101x* was *climate change* - the third most frequently-used bigram. The analysis also suggests that students taking the course were interested in geology. The word, *geology,* was among the first 20 most frequently-used words (n=76), and the analysis identified other frequently-used language chunks (*earth science/s* and *interest in geology*) related to the topic. To confirm this deduction, we used the search option in the worksheet view to examine in what contexts the word, *geology,* was used (Figure 2). This examination showed that students were, indeed, interested in learning about the geology of Antarctica; moreover, many of them already had solid backgrounds in geology, including undergraduate and graduate qualifications, and some were geology teachers.



Response	Words	Sentences
Hello I'm [redacted]. I'm from Missouri in the US and I teach high school science. I'm talking this class to further my education in geology and find new exciting content to share with my students.	39	3
Hello everyone, my name is [redacted] and I'm a textbook writer and editor. I live in Boynton Beach, Florida, and I'm taking the course because I'm interested in geology and seeing places i've never been.	39	2
Hi, I'm from Guatemala. I'm interested in geology since I can remember. I don't really have a direct connection with Antarctica but I really would like to be part of an study there. Good luck everyone!	39	4
Hello! My name is [redacted] I am a recent university graduate in geological engineering. I live just outside of Toronto and am taking this course because I have a passion for both geology and history!	36	3
Hi everybody! My name is [redacted]. IÂ´m Geologist and IÂ´m from Colombia. I am keen to know more about AntarticaÂ´s geology and evolution as well as all the implications it has with global change. See you!	36	4
Hi everyone I'm [redacted] I study Antarctica at university from a climate change perspective but its been a while since I did geology so I'm hoping for a refresher from this course!	34	1
Hi! My background is in human physiology, biochemistry and nutrition. Intrigued with geology, especially after touring and learning about the Grand Tetons and Yellowstone Park in Wyoming USA. Plate tectonics fascinate me.	32	3
I'm from South Brazil living in Rio de Janeiro. I just finished my PhD in analysis of Antarctic Ice core samples, and I would like to learn about Antarctic geology.	31	2
Hi, I'm [redacted] living in London, UK. I've read so much about Antarctic exploration - real and fictional and am looking forward to finding out more, especially the geology.	30	2
Hi. I'm [redacted] from Wellington, New Zealand. I visited Antarctica on holiday in January and am looking forward to finding out more about its history and geology.	28	2

Figure 2: Quantext analysis: A worksheet view

Although *history* was not among the 30 most frequent words, we also searched for it in the worksheet view because historical thinking was another important aspect of the course. The search returned 50 hits, which showed that about 11% of the students acknowledged history of Antarctic exploration as a content area of interest. However, they did not show prior knowledge of this topic.

The basic Quantext analysis also revealed a theme of personal learning goals. *Course*, was the second most frequently-used word (n=253) and *learn* was the fourth (n=141), with *learning* (n=88), *know* (n=62) and *knowledge* (n=53) among the 30 most frequent words. The bigram and trigram analyses showed both general interest in learning (*forward to learning*; *like to learn*; *learn new things*; *keen to learn*; *hope to learn*; *like to know*; *broaden my knowledge*; *love to learn*) and a more specific interest in geology, earth and environmental sciences. High-frequency bigrams displayed students' enthusiasm about the course (*look/ing forward*; *learn new*; *new things*; *'d like*; *'d love*). Other frequent vocabulary was indicative of sharing personal details (*currently living*; *years ago*; *years old*; *high school*; *long time*). *New Zealand* was the second most frequently-used bigram and *United Sates* the sixth. The social function of the introductory forum was apparent in the frequent use of greetings; the words *Hi* and *Hello* were the third and fifth most frequently-used in the forum.

In summary, the first finding of this analysis is that students followed the suggested topics in their introductory posts. It is notable that climate change was shown to be a key area of students' interest in this course. Knowing this, we now move to the second part of the analysis – students' posts in topic related course discussion forums.

3.2 A discussion forums' case

We found that, in topic-related discussions, students actively engaged with the prompts and questions related to the topic of climate change, which they identified as a key point of interest in the introductory forum. Although the overall number of posts in the discussion forums reduced gradually throughout the course, and was paralleled by active student numbers in the course by week, an examination of the descriptive statistic for the number of words per post showed a different picture.

The longest posts were in response to q10 (M = 93), q12 (M = 118) and q13 (M = 89) in week 4. Interestingly, these questions are all related to climate change. *Climate change* was the most frequent bigram in students' responses to q12 and q13, and among the first 10 bigrams for q10. Other frequent bigrams across the three questions were: *global warming*, *ice sheet/s*, *global temperature*, *greenhouse gases*, *climate system*, *carbon dioxide*, *fossil fuel/s*, *sea level*, *CO2 emissions*, and *renewable energy;* all of which are closely related to the topic of climate change. The fourth longest post average (M = 87 words) was recorded in week 2 for q6, also related to the broader topic of climate change. Here, the most commonly-used bigrams were: *ozone depletion*, *UV radiation*, *ozone hole*, *ozone layer;* while *ozone depletion impact* was among most frequent trigrams. An additional analysis showed that the words, *climate* and *ozone,* had high keyness values (1162.045 and 957.496, respectively) across posts in all course forums, compared to the frequency of these words in general academic texts (in the COCA corpus), further confirming their prominence in the students' discourse. Responses to q10 and q12 were also characterised by the highest mean lexical density (.61-.62), a feature of academic writing indicating higher complexity.

The shortest responses (in words per post) were to q14 - week 5 (M = 54), q2 - week 1 (M = 60), q11 - week 4 (M = 62) and q4 - week 2 (M = 63). These questions focused on creating opportunities for students to share personal experiences, e.g., their favourite art or books about Antarctica, their interest in visiting Antarctica, or personal opinions about preservation of animals and birds in Antarctica. Ideas expressed by students' in q2 and q14 were also less complex, as indicated by their lower mean lexical density.

4 IMPLICATIONS AND FUTURE WORK

Based on the results of our proof of concept study, we believe there is merit in using Quantext as a way of interrogating students' responses to open-ended questions in MOOC discussion forums. This is because even the basic text analysis (as the one we conducted) may offer MOOC instructors insights into students' interests in and engagement with course topics, at different points in the course. Our results suggest that instructors would do well to analyse introductory students' posts for main topics and areas of interest, and to adjust course discussion prompts to reflect these interests. This approach might lead to deeper thinking and higher levels of engagement in MOOC discussions.

The present study reports only on the basic text analysis of student posts in a MOOC. In our future analyses, we plan to (1) compare patterns of language use in student posts and reference corpora of the MOOC learning materials, (2) use automatically generated similarity indices to compare student responses to prompts with model answers, (3) explore the effect of question type on the nature of student responses, and (4) investigate links between students' contributions to MOOC discussions and their course assessment data.

REFERENCES

Balakrishnan, G. & Coetzee, D. (2013). *Predicting student retention in Massive Open Online Courses using Hidden Markov Models*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.659.2890&rep=rep1&type=pdf

Beheshtiha, S. S., Hatala, M., Gašević, D., & Joksimović, S. (2016). The role of achievement goal orientations when studying effect of learning analytics visualizations. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 54-63). Edinburgh, United Kingdom: The University of Edinburgh.

Boyer, S., & Veeramachaneni, K. (2015). Transfer learning for predictive models in Massive Open Online Courses. In C. Conati, N. Heffernan, A. Mitrovic, M. Verdejo (Eds.) *Artificial Intelligence in Education*. AIED 2015. Lecture Notes in Computer Science, vol. 9112. Springer, Cham.

Crossley, S., Paquette, L., Dascalu, M., McNamara, D., & Baker, R. (2016). Combining click-stream data with NLP tools to better understand MOOC completion. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 6-14). Edinburgh, United Kingdom: The University of Edinburgh.

Jiang, S., Williams, A. E., Schenke, K., Warschauer, M., & O'Dowd, D. K. (2014). Predicting MOOC performance with Week 1 Behavior. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.) *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 273-275). London, United Kingdom: Institute of Education, UCL.

Jimoyiannis, A., & Angelaina, S. (2012). Towards an analysis framework for investigating students' engagement and learning in educational blogs. *Journal of Computer Assisted Learning, 28*, 222–234.

Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., and Friss de Kereki, I. (2016). Translating network position into performance: Importance of centrality in different network configurations. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 314-323). Edinburgh, United Kingdom: The University of Edinburgh.

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., & Siemens, G. (2016). Towards automated content analysis of discussion transcripts: A cognitive presence case. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 15-24). Edinburgh, United Kingdom: The University of Edinburgh.

McDonald, J. & Moskal A.C.M. (2017) Quantext: analysing student responses to short-answer questions. In H. Partridge (Ed.) *Me, Us, IT! Proceedings ASCILITE2017: 34th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education*. Toowoomba, Australia: University of Southern Queensland.

McNamara, D. S., Allen, L. K., Crossley, S. A., Dascalu, M., & Perret, C. A. (2017). Natural language processing and learning analytics. In G. Siemens & C. Lang (Eds.) *Handbook of learning analytics and educational data mining* (pp. 93-104). Society for Learning Analytics Research (SOLAR).

Piech, C., Huang, J., Chen, Z., Do, C. B., Ng, A. Y., Koller, D. 2013. Tuned models of peer assessment in MOOCs. S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.) In *Proceedings of The 6th International Conference on Educational Data Mining*. Memphis, TN: FedEx Institute of Technology.

Poquet, O., & Dawson, S. (2016). Untangling MOOC Learner Networks. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 208-212). Edinburgh, United Kingdom: The University of Edinburgh.

Reich, J., Tingley, D., Lede-Luis, J., Roberts, M. E., & Stewart, B. M. 2015. Computer-assisted reading and discovery for student-generated text in Massive Open Online Courses. *Journal of Learning Analytics, 2*(1), 156–184.

Sharma, K., Jermann, P., & Dillenbourg, P. (2015). Identifying styles and paths toward success in MOOCs. In O. C. Santos, J. G. Boticario, C. Romero, M. Pechenizkiy, A. Merceron, P. Mitros, J. M. Luna, C. Mihaescu, P. Moreno, A. Hershkovitz, S. Ventura, & M. Desmarais (Eds.) *Proceedings of the 8th International Conference on Educational Data Mining*. Madrin, Spain: Spanish National University for Distance Education.

Wang, X., Wen, M., & Rose, C. 2016. Towards triggering higher-order thinking behaviors in MOOCs. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 398-407). Edinburgh, United Kingdom: The University of Edinburgh.

Wen, M., Yang, D., & Rosé, C. P. 2014. *Sentiment analysis in MOOC discussion forums: What does it tell us?* In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.) *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 130-137). London, United Kingdom: Institute of Education, UCL.

Studying Adaptive Learning Efficacy using Propensity Score Matching

Shirin Mojarad[1], Alfred Essa[1], Shahin Mojarad[1], Ryan S. Baker[2]
McGraw-Hill Education[1], University of Pennsylvania[2]
{shirin.mojarad, Alfred.essa, s.a.mojarad}@mheducation.com, Rybaker@upenn.com

ABSTRACT: Many higher education institutions are in the process of adopting adaptive learning platforms for online and hybrid learning. However, it is unclear how effective these platforms are at improving student success rates. ALEKS (Assessment and LEarning in Knowledge Spaces) is an adaptive learning system designed for courses in science and mathematics. ALEKS has several mathematics courses that cover developmental mathematics for both four year and two year colleges. In this study, we investigate the effectiveness of ALEKS at a community college, where some sections have adopted ALEKS while others chose not to use it. We conduct different possible comparisons of ALEKS versus Non-ALEKS sections and students, including conducting a quasi-experiment using propensity score matching (PSM) to construct two similar groups of learners to compare between. PSM is conducted by matching ALEKS and Non-ALEKS users across their Accuplacer score, age and race. In all comparisons, students using ALEKS have significantly higher pass rates than comparison groups. When matching students using PSM, students who use ALEKS pass 19 percentage points more often than students who do not use ALEKS.

Keywords: efficacy, adaptive learning, ALEKS, propensity score matching, quasi-experimental design.

1 INTRODUCTION

Adaptive learning is increasingly being adopted across different institutions and disciplines to improve student outcome (Kolb & Kolb, 2005). However, it remains unclear how effective many of these platforms are at producing positive outcomes in these settings. It is important to investigate the efficacy of adaptive platforms in different educational settings, to help instructors, institutions, and students to decide which platforms to use in their classes (Hallahan, Keller, McKinney, Lloyd, & Bryan, 1988).

In this study, we examine the effectiveness of ALEKS (Assessment and LEarning in Knowledge Spaces), a widely used adaptive learning system, in the context of a community college. Although there is evidence for ALEKS's effectiveness in other contexts, including in K-12 schools (Craig et al., 2013) and in non-traditional adult learning settings (Rivera, Davis, Feldman, & Rachkowski, 2017), there is not yet solid evidence on the efficacy of ALEKS in a community college setting.

We investigate this question within the context of a large community college in the Midwestern United States. According to many, a randomized controlled trial (RCT) would be the gold standard, ideal way to investigate this question (Silverman, 2009). However, RCTs are costly, and sometimes not feasible to conduct in educational settings, both due to difficulty in securing agreement for randomized assignment, and due to challenges in establishing implementation fidelity (Feng, Roschelle, Heffernan, Fairman, & Murphy, 2014). For these reasons, many researchers have argued for the use of quasi-experiment studies besides or instead of RCTs. In these studies, subjects are assigned to treatment and control groups based on some criteria such as subjects' date of birth, while in RCTs this assignment is random. Quasi-experiment studies are a practical and acceptable alternative to RCTs when design of an RCT is implausible (Sullivan, 2011) .

Within this specific community college, it was not practical to randomly assign instructors or classes to conditions, as the college has made a policy decision that eliminates the ability to use an RCT design to study the efficacy of its chosen product. Instead, the college's administration decided that instructors would be given the choice of adopting ALEKS in their courses, and many instructors chose not to use it. Even when an instructor did choose to adopt ALEKS, it was not required for students, and was counted minimally towards the final grade. Therefore, only a portion of students in classes adopting ALEKS ever used ALEKS, and many students may not have used ALEKS to the degree or in the fashion intended. Therefore, we probably should not simply compare ALEKS classes to non-ALEKS classes; there are both selection bias issues and valid concerns about implementation fidelity (Feng et al., 2014).

An alternative would be to simply compare between the students who used and did not use ALEKS, ignoring what classes they were in. However, since this study was not designed as an RCT, there are issues of selection bias in making this comparison – it is possible, for example, that the students who decided to use ALEKS could have been the strong students to begin with and would have done well in the course anyways.

Therefore, in addition to investigating the effectiveness of ALEKS by comparing between different naturally-occurring student populations, we design a quasi-experiment study to Isolate the effects of student characteristics and find comparable student populations who mainly differ in their use of ALEKS (P. R. Rosenbaum, 2010). This is done using propensity score matching (PSM) (Austin, 2011). Propensity score matching is explained in more detail in section 2. Section 3 explains the data used in this study and the study design. Sections 4 and 5 cover the results and conclusions of the study.

2 PROPENSITY SCORE MATCHING

In RCT, random allocation is used to choose the treatment and control groups, so that study subjects have the same chance of being assigned to each study group. However, as described in section 1, random assignment is not plausible in many studies. For example, in the current study, learners are assigned to the ALEKS or non-ALEKS group depending on the class they have registered for and whether they chose to use ALEKS. The methods to study such groups are often described as quasi-experimental (Cochran & Chambers, 1965). The main concern in these studies is selection bias, since subjects are assigned to each group based on a criteria, and therefore they might not have similar chances of being assigned to treatment and control groups. PSM is a method that is used to remove this bias by finding control and treatment groups from the study cohort, such that they have similar probability of being assigned to the control and treatment group, at least according to a set of "baseline characteristic" variables describing members of the population. Therefore, it creates a study that resembles an RCT.

Considering two possible outcomes of receiving and not receiving treatment, each learner has two potential outcomes of $Y_i(0)$ and $Y_i(1)$, the outcomes under the control and treatment, respectively. However, each learner is either in the control or treatment group. We define Z as an indicator variable on whether the learner received the treatment (Z = 0 for control/non-ALEKS vs. Z = 1 for treatment/ALEKS). Therefore, we can only observe one outcome for each learner.

In PSM, for each member of the intervention group, we identify a member of the control group that is as similar as possible in terms of their propensity score. Then, the difference in outcomes between the matched pair is computed. The average of this difference over the observed pairs is an estimate of the mean causal effect of a particular intervention on outcome.

A propensity score is used to choose treatment and control groups with similar baseline characteristics. A propensity score is defined as the probability of the subjects being assigned to the treatment group, given a set of baseline characteristics (P. Rosenbaum & Rubin, 1983). This can be formulated as the conditional probability of being exposed to intervention given baseline characteristics X:

$$e_i = P\,(Z_i = 1\,|\,X_i)$$

where e_i is the propensity score and X is the vector of observed characteristics of the subject. This can be modelled as a logistic regression model, where the dependent variable is the probability of receiving treatment and independent variables are the baseline characteristics:

$$p(Z = 1|X) = \frac{1}{1 + e^{-\beta X}}$$

One advantage of PSM is that the regression model used to predict the probability of receiving treatment takes into account the relationship between baseline characteristics. In addition, PSM enables matching not just at the mean but balances the distribution of observed characteristics across treatment and control.

3 DATA AND METHODOLOGY

3.1 Data

We obtained data from 3422 students in 198 sections covering four courses including pre-algebra (67 sections), elementary algebra (44 sections), intermediate algebra (43 sections) and college math (44 sections). Amongst these, 37 sections with 706 students adopted ALEKS. From these students, only 417 (59%) used ALEKS. Figure 1 shows a representation of ALEKS and non-ALEKS sections and students.

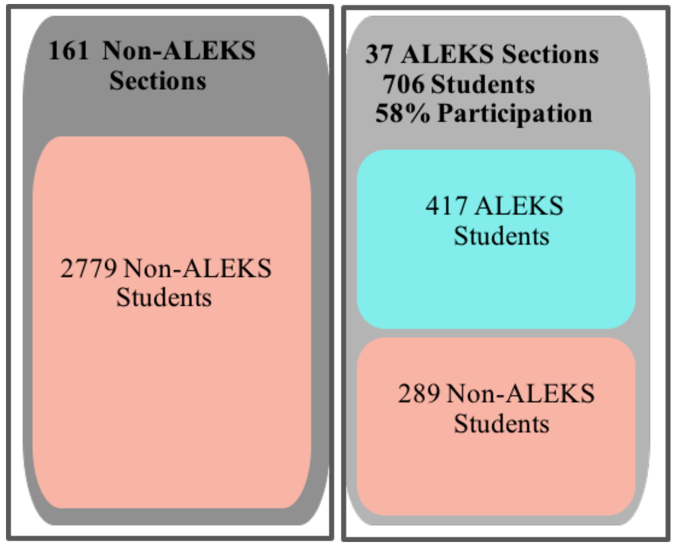


Figure 1: Breakdown of ALEKS and non-ALEKS sections and students.

3.2 Methodology

We have made comparisons of several possible breakdowns of ALEKS vs. Non-ALEKS students and sections. Below are the comparisons conducted in this study:

1. ALEKS students vs. all Non-ALEKS students (in both ALEKS and non-ALEKS sections)
2. ALEKS students in ALEKS sections vs. Non-ALEKS students in ALEKS sections
3. ALEKS students in ALEKS sections vs. Non-ALEKS students in Non-ALEKS sections
4. ALEKS sections vs. Non-ALEKS sections
5. Matched ALEKS students vs. Matched Non-ALEKS students

What could differentiate students in comparison groups 1-4 is their starting knowledge and their current learning situation which could be affected by their age and race. Therefore, the last comparison is done by matching ALEKS and non-ALEKS users using PSM. Matching is done using three student characteristics: Accuplacer arithmetic score, age, and whether the student's race is classified as minority or not. The Accuplacer score is used by the college to decide whether to place students into developmental math courses and is used as a measure of students' prior knowledge in the subject (Mattern & Packman, 2009). The student group from which the control matches where identified includes only non-ALEKS students in non-ALEKS sections. This naturally removes the student selection bias in the control group, since students in non-ALEKS sections do not have a choice to use ALEKS.

A logistic regression model is used to calculate the propensity score of students -- specifically, the binomial generalized linear model from *statsmodels* package in Python was used. The logistic regression model had whether the student used ALEKS as a binary outcome and the independent attributes consisted of Accuplacer arithmetic score, age, and whether the student race is minority or not. Figures 2.a-2.d shows the distribution of each of these attributes and the propensity score of ALEKS and non-ALEKS users before and after matching. Matching on propensity score is conducted as a 1-1 matching using nearest neighbor approach, which uses the distance between propensity scores to find the closest match. Hence, for each treatment subject, a control match is selected as the subjects with the closest propensity score.

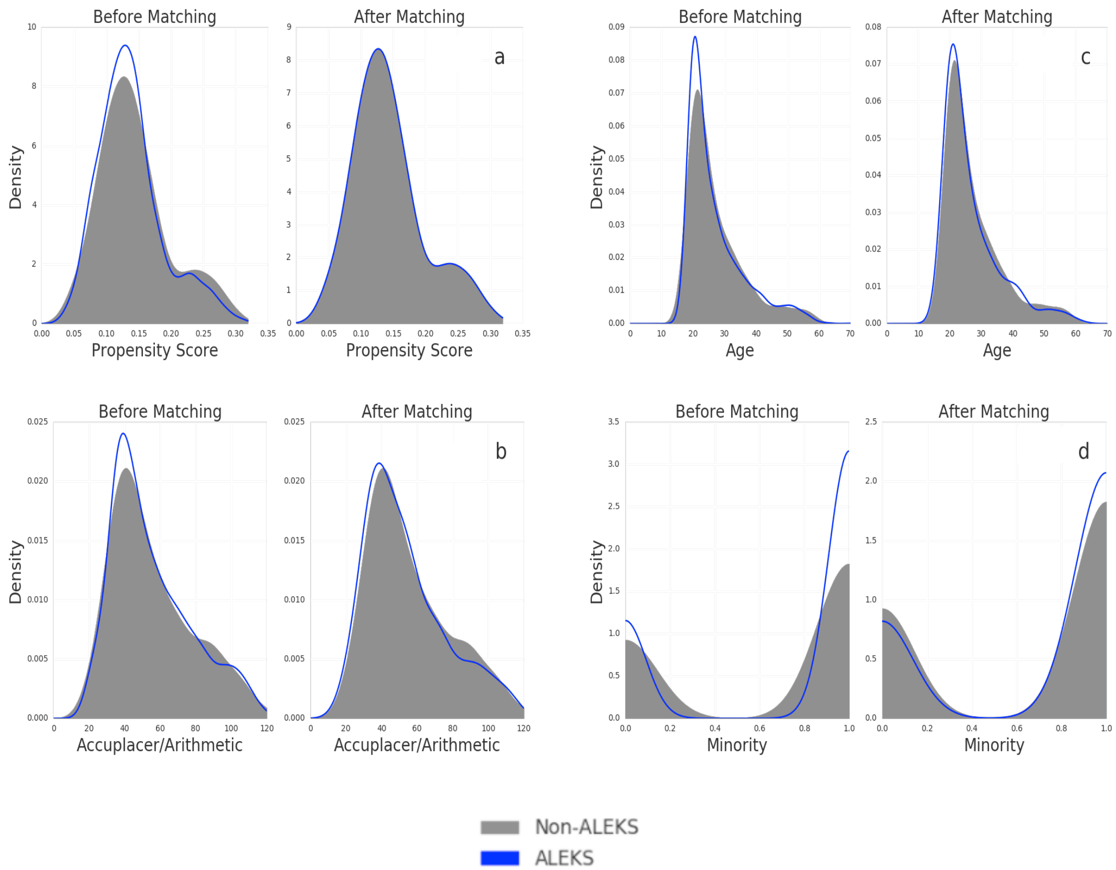


Figure 2: distribution of a) propensity score b) Accuplacer score c) age, d) minority, before and after matching for ALEKS (blue) and non-ALEKS (grey) users.

4 RESULTS

Table 1 shows the ALEKS and non-ALEKS group pass rates, the ALEKS-non-ALEKS group difference in pass rates , and p-value for each of the comparisons mentioned in section 3.2. The criteria for pass is grades C+ and above. Students' grades are measured using a uniform test conducted across all classes at the end of semester. We have used the chi-square (χ^2) contingency test to compare the pass rate across two groups (Rao & Scott, n.d.). We conduct five comparisons. The first comparison is all students who at least took an initial assessment in ALEKS (ALEKS students) versus all students who did not use ALEKS in the course of the class (non-ALEKS students). Within this comparison, ALEKS students had statistically significantly higher pass rates, χ^2(df=1, N=3422) = 29.5, p<0.001, with ALEKS achieving a boost of 14 points in pass rates.

We have used the chi-square (χ^2) contingency test to compare the pass rate across two groups (Rao & Scott, n.d.). We conduct five comparisons. The first comparison is all students who at least took an initial assessment in ALEKS (ALEKS students) versus all students who did not use ALEKS in the course of the class (non-ALEKS students). Within this comparison, ALEKS students had statistically significantly higher pass rates, χ^2(df=1, N=3422) = 29.5, p<0.001, with ALEKS achieving a boost of 14 points in pass rates.

The second comparison is between ALEKS and non-ALEKS only within ALEKS sections. This comparison is important as it naturally controls for the instructor and class environment, by comparing students who did and did not use ALEKS within the same class. Within this comparison, ALEKS students had statistically significantly higher pass rates, χ^2(df=1, N=709) = 24.7, p<0.001, with ALEKS achieving a boost of 19 points in pass rates.

The third comparison considers assignment at the classroom level. In this comparison, all students within ALEKS sections, whether they did or did not use ALEKS, are compared against all the students in non-ALEKS sections. This comparison is perhaps the most standard quasi-experimental comparison, but raises questions of implementation fidelity. Within this comparison, ALEKS students had statistically significantly higher pass rates, χ^2(df=1, N=198) = 8.1, p=0.004, with ALEKS achieving a boost of 6 points in pass rates.

The fourth comparison is between ALEKS students in ALEKS sections and non-ALEKS students in non-ALEKS sections. Within this comparison, we are excluding non-ALEKS students in ALEKS sections from this comparison as those are the students who chose not to use ALEKS, despite having the option of using it in the class. Including these students includes students who did not participate in the treatment, despite being assigned to the treatment group, creating questions of implementation fidelity. Within this comparison, ALEKS students had statistically significantly higher pass rates, χ^2(df=1, N=3196) = 27.5, p<0.001, with ALEKS achieving a boost of 14 points in pass rates.

Finally, comparison five attempts to avoid the biases inherent in the first four comparisons, comparing ALEKS students who are matched with similar non-ALEKS students in non-ALEKS classes. The matching is done using Accuplacer, age and minority and as shown above, the students selected in the matching process have similar prior knowledge, age, and minority between conditions. All students in the matched treatment condition used ALEKS and all students in the matched control condition did not use ALEKS. Within this comparison, ALEKS students had statistically significantly higher pass rates, χ^2 (df=1, N=748) = 16.5, p<0.001, with ALEKS achieving a boost of 15 points in pass rates.

As shown in this table, all comparisons are statistically significantly in favor of ALEKS, with a boost of 6 to 19 points in pass rates between ALEKS and non-ALEKS users across different comparisons. Some of the comparisons are likely to be biased in favor of ALEKS, others against ALEKS, but overall they tell a common story – ALEKS is statistically significantly more effective at enhancing pass rates compared to the control condition.

Table 1: Pass rates and significance level for ALEKS and non-ALEKS users.

Comparison	Pass Rates for ALEKS vs. Non-ALEKS	Boost	p-value
1. ALEKS students vs. all Non-ALEKS students	71% vs 57%	+14	<0.001
2. ALEKS students vs. Non-ALEKS students in ALEKS sections	71% vs 52%	+19	<0.001
3. ALEKS sections vs. Non-ALEKS sections	63% vs 57%	+6	0.004
4. ALEKS students in ALEKS sections vs. Non-ALEKS students in Non-ALEKS sections	71% vs 57%	+14	<0.001
5.Matched ALEKS students vs. Matched Non-ALEKS students (quasi-experiment study using Propensity Score Matching)	70% vs 55%	+15	<0.001

5 CONCLUSION

In this paper, we present a study on the pass/fail outcomes of students who used and did not use ALEKS within their developmental math courses. For several comparisons, the results show significantly higher pass rates amongst students using ALEKS. However, it may not be valid to compare between groups directly, due to concerns around selection bias and implementation fidelity. Therefore, we conducted a quasi-experiment study using Propensity Score Matching (PSM), labeled comparison 5 in Table 1. The results show that students using ALEKS have significantly higher pass rates, even when we use PSM to control for students' math placement score, age and race. However, as with all PSM-based quasi-experiment study, it may be that the matching was imperfect. It is important to consider other factors that could affect student outcome besides the ones used for matching ALEKS and non-ALEKS users in this study. Other factors could include but are not limited to social-economic background, and prior academic performance, as well as students' attitudes towards learning and attitudes towards online learning technologies. As such, conducting further follow-up studies will help us more conclusively understand whether ALEKS is positively benefitting students.

Following the results of this study, the college in which we conducted the study is in process of adopting ALEKS in more sections, and encouraging instructors to make ALEKS a requirement for students and a part of their grades. We intend to follow-up this study with a subsequent study, at the same institution, to see if usage has genuinely increased, and if so, whether the greater proportion of ALEKS users maintain the same improvements in outcomes seen in this study. Within this upcoming study, we intend to also control for a greater range of factors. By doing so, we may be able to better understand the degree to which ALEKS is benefitting students, and whether these benefits are equivalent across all groups of students.

ACKNOWLEDGMENTS
[Redacted for submission]

REFERENCES

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of

Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Cochran, W. G., & Chambers, S. P. (1965). The Planning of Observational Studies of Human Populations. *Journal of the Royal Statistical Society. Series A (General)*, *128*(2), 234. https://doi.org/10.2307/2344179

Craig, S. D., Hu, X., Graesser, A. C., Bargagliotti, A. E., Sterbinsky, A., Cheney, K. R., & Okwumabua, T. (2013). The impact of a technology-based mathematics after-school program using ALEKS on student's knowledge and behaviors. *Computers and Education*, *68*(October), 495–504. https://doi.org/10.1016/j.compedu.2013.06.010

Feng, M., Roschelle, J., Heffernan, N., Fairman, J., & Murphy, R. (2014). Implementation of an Intelligent Tutoring System for Online Homework Support in an Efficacy Trial (pp. 561–566). Springer, Cham. https://doi.org/10.1007/978-3-319-07221-0_71

Hallahan, D. P., Keller, C. E., McKinney, J. D., Lloyd, J. W., & Bryan, T. (1988). Examining the Research Base of the Regular Education Initiative. *Journal of Learning Disabilities*, *21*(1), 29–35. https://doi.org/10.1177/002221948802100106

Kolb, A. Y., & Kolb, D. A. (2005). Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. *Academy of Management Learning & Education*, *4*(2), 193–212. https://doi.org/10.5465/AMLE.2005.17268566

Mattern, K. D., & Packman, S. (2009, December 4). Predictive Validity of ACCUPLACER Scores for Course Placement: A Meta-Analysis. Research report, The College Board. Retrieved from https://research.collegeboard.org/publications/content/2012/05/predictive-validity-accuplacer-scores-course-placement-meta-analysis

Rao, J. N. K., & Scott, A. J. (n.d.). On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated from Survey Data. *The Annals of Statistics*. Institute of Mathematical Statistics. https://doi.org/10.2307/2241033

Rivera, M. A., Davis, M. H., Feldman, A., & Rachkowski, C. (2017). An outcome evaluation of an adult education and postsecondary alignment program: the Accelerate New Mexico experience. *Problems and Perspectives in Management (Open-Access)*, *11*(4). Retrieved from https://businessperspectives.org/journals/problems-and-perspectives-in-management/issue-40/an-outcome-evaluation-of-an-adult-education-and-postsecondary-alignment-program-the-accelerate-new-mexico-experience

Rosenbaum, P. R. (2010). *Design of Observational Studies*. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-1213-8

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*, 41–55. https://doi.org/10.1093/biomet/70.1.41

Sullivan, G. M. (2011). Getting Off the "Gold Standard": Randomized Controlled Trials and Education Research. *Journal of Graduate Medical Education*, *3*(3), 285–289. https://doi.org/10.4300/JGME-D-11-00147.1

Beyond Automated Essay Scoring: Forecasting and Improving Outcomes in Middle and High School Writing

Elijah Mayfield, David Adamson, Bronwyn Woods,
Shayne Miel, Stephanie Butler, and Jill Crivelli
Turnitin, 2020 Smallman St, Pittsburgh PA 15232
{emayfield, dadamson, bwoods, smiel, sbutler, jcrivelli}@turnitin.com

ABSTRACT: This paper presents an analysis of an automated essay scoring (AES) system in two studies of live classroom use. First, in a study of 99 students in Texas, we show that automated scores do predict future performance on standardized tests, and that in-system activity can be included in a predictive model to further improve accuracy. Following that, the results of a five-school study in Maryland demonstrate moderate evidence that automated essay scoring is correlated with school-level improvement in ELA outcomes.

Keywords: Automated essay scoring, intelligent tutoring systems, writing assessment, school implementation, quasi-experimental efficacy

1 INTRODUCTION

Student performance in writing is difficult to assess at large scales, and targeted instruction based on that assessment is even more challenging. Unlike in math or reading, turnaround time for even short written student work can take weeks, and large-scale assessment for schools or districts may not be available until the following school year. Scoring relies on instructors or trained scorers who can become tired or distracted over hours of scoring, leading to inconsistent results (Williamson et al., 2012). English Language Arts (ELA) teachers at these grade levels can also teach up to 6 classrooms, up to 200 students at a time, which combined with low salaries and minimal support leads to particularly high attrition, low job satisfaction, and poor student outcomes (Scherff & Hahs-Vaughn, 2008).

Automated essay scoring (AES) aims to solve some of these problems. For half a century, researchers have have worked to reduce the time burden of (Page, 1966). This goal remains largely consistent today. AES models are trained on a small set of essays scored by hand, and then score new essays with the reliability of an expert rater. A large body of work, particularly in the last decade, has demonstrated this reliability (Shermis & Burstein, 2013).

This paper investigates AES in classrooms over time, evaluating the relationship of AES use to outcomes in two authentic settings. The first section studies the use of AES for *predicting* outcomes of individual students in a Texas high school. The second section evaluates whether AES *improves* ELA outcomes in several Maryland middle schools. This second causal claim is a much more challenging bar for AES than scoring accuracy or forecasting ability.

2 BACKGROUND

AES has focused historically on replicating expert readers for large-scale scoring of thousands of essays, either for end-of-year standardized assessments or entrance exams like the GRE or TOEFL (Attali & Burstein, 2004). This use preferences interpretable model features informed by psychometrics, often representing high-level characteristics of writing like coherence or lexical sophistication. The primary goal is defensibility of the underlying model, known as construct validity. This construct validity through feature choice has been emphasized over measuring the ability to provide actionable guidance to writers based on the scoring.

In the 1990s and early 2000s, classroom technology was released based on this approach, including ETS Criterion, Pearson WriteToLearn, and Vantage MyAccess. Classroom reviews of these products were mixed at best. While their use positively impacted student writing (Shermis et al., 2008), students felt negative about the experience (Scharber et al., 2008). The most widely-cited districtwide study on these tools (Grimes & Warschauer, 2010) described the work as "fallible" and gains in school outcomes were not demonstrated. Teachers using earlier tools stated that automated scoring must be paired with actionable next steps for writers (Riedel et al., 2006). Building on this, work in academic settings has used AES to provide formative writing instruction and feedback that students perceive as "informative, valuable, and enjoyable" (Roscoe et al., 2013) and which provides more efficient learning gains than practice alone (Crossley et al., 2013).

Alongside the emergence of that research, a newer generation of tools has refocused AES to prioritize feedback to students. These include TenMarks Writing, WriteLab, Grammarly, PEG Writing, and Turnitin Revision Assistant. AES feedback's impact on writing quality varies by product. For instance, PEG Writing has been shown to save teachers time and let them focus on higher-level writing skills, but not to improve writing quality (Wilson & Czik, 2016). Revision Assistant provides feedback that students rate as helpful, and encourages editing that improves quality across drafts (Woods et al., 2017). To date, there is little work discussing the longitudinal effect of AES on classroom instruction during the school year.

2.1 Turnitin Revision Assistant

This research focuses on two school districts and one AES technology used in both, Turnitin Revision Assistant, released in 2016. Prior work has shown that the AES used in this product reliably predicts student writing scores in line with the state-of-the-art (Shermis, 2014).

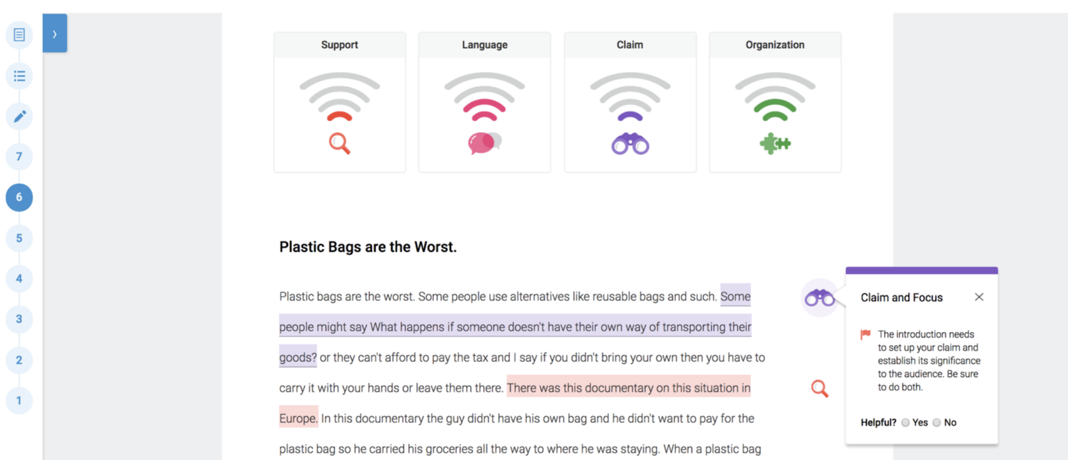


Figure 1: Automated essay scoring through Signal Checks in Revision Assistant.

In Revision Assistant, students request feedback from a “Signal Check”. This provides automated scores on rubric traits in a visual format (Figure 1) and highlights up to four sentences within the text for in-line feedback. The full feedback algorithm is described in Woods et al (2017). Revision Assistant also contains “Spot Check” assessments, which remove real-time feedback, instead scoring essays for teacher review. This Spot Check environment matches summative settings like standardized testing and gives teachers insight into student skill transfer into settings where real-time support will not be available. Note that in this paper, a “draft” of student work corresponds to a Signal Check or a final submission; intermediate work between requests for feedback is not a separate draft.

Based on prior literature, AES feedback in Revision Assistant should have a positive impact on classrooms. Students who learn to think of writing as a process that includes iterative improvement demonstrate large gains in transferable skills (Dix, 2006; Tillema et al., 2011). Unfortunately, this process is difficult to learn and complex to teach, needing differentiated instruction across students and incorporating strategies that may vary across tasks (Hayes & Flower, 1980). Teachers tend to view this element of instruction as difficult and time-consuming, and rarely teach the revision process in depth (Graham & Harris, 2005).

3 FORECASTING STUDENT OUTCOMES

In Texas, student performance is evaluated on the State of Texas Assessments of Academic Readiness, or STAAR (Texas Education Agency, 2017). This test measures student progress against curricula aligned to Texas Essential Knowledge and Skills, or TEKS, standards. Students in grades 3-11 are assigned a Reading component, while writing is evaluated in the 4th, 7th, and 9th-11th grades. Writing Scores are broken out separately and are also combined with Reading scores into an overall ELA Score. This study evaluates the use of Revision Assistant to forecast student outcomes on the STAAR assessments on both the Writing Score and the combined ELA Score. We find that the Revision Assistant forecast compares favorably to, and effectively supplements the information provided by, the existing Fall benchmark currently used by the school.

3.1 Methods

Six English I (9th grade) classes from a large, urban school district, taught by four teachers, were selected to participate in the study. In January, teachers met and were trained on the AES system, including the difference between Signal Check and Spot Check assignments. A total of 111 students were enrolled in participating classes during the administration of a school-wide benchmark in fall 2017; of those, 99 participated in the study. Shortly after training, teachers administered an initial, timed Spot Check assignment. Teachers were then given access to Revision Assistant for three months, with a recommended pacing guide that included four writing prompts appropriate to the school curriculum and sequencing of English I. Adherence to this pacing guide was not mandated. A second Spot Check assessment was administered, no more than one month prior to the STAAR assessment. Spot Check assignments matched the genre of writing used in the STAAR assessment, though there were some differences. For instance, the writing was typed instead of handwritten, and was not subject to length constraints (Texas students are penalized for exceeding a maximum length on standardized essay assessments). End-of-year STAAR testing, administered at the end of March 2017, was used for final performance evaluation.

We first evaluate the pre-existing benchmark and the results from Spot Check assessments as individual, linear predictors of STAAR performance. Next, we fit a multivariate linear regression using four variables and use that regression to predict overall STAAR ELA and STAAR Writing performance. In this model, the first three variables are direct student evaluations: the **pre-existing benchmark score**, student performance on the **initial Spot Check**, and student performance on the **second Spot Check**. The fourth is a measure from Signal Check assignments during the class curriculum - specifically, the **total count of Invalid Drafts** submitted by each student. An Invalid Draft is a draft that was not given a score, due to being off-topic or in bad-faith. Detection of such drafts is fully automated through machine learning. Invalid drafts can represent student "churn" - an inability to compose essays that meet assignment criteria - or student disengagement. Both are early warning signs that can be addressed through targeted instruction.

We evaluated other factors from formative assignments in a Signal Check setting, such as total number of drafts authored, growth (as measured by automated scoring) within-assignment, and increases in word count. In a forward stepwise regression, after including variables for student performance on Spot Check assignments, none of these other factors improved model fit or were significant in a *t*-test. They are not included in our results.

3.2 Results

The school district's existing fall benchmark is reasonably reliable at forecasting student performance on the STAAR ELA assessment as a whole (r = 0.63). However, it is only slightly predictive of student end-of-year STAAR Writing performance (r = 0.26). As this has historically been the only available benchmark, the school has had no access to actionable insights on student writing performance. The first Spot Check, by contrast, was moderately predictive both of STAAR Writing (r = 0.45) and overall STAAR ELA (r = 0.43) performance.

Table 1: Accuracy of models by correlation (r) and root mean squared error (RMSE)

	STAAR Writing (2-8 scale)		**STAAR ELA (0-68 scale)**	
	r	RMSE	r	RMSE
Existing Benchmark	0.26	0.97	0.63	6.63
Spot Check Assessment	0.45	0.90	0.43	7.70
4-Variable Forecast	0.58	0.82	0.74	5.42

In the multivariate model, the addition of AES variables significantly increases predictive accuracy over the benchmark alone, which combine to explain 55% of student performance (r^2) on STAAR ELA. The four variables are all significant ($p < 0.01$) in both forecasting models. This analysis is summarized in Table 1; scatter plots are presented in Figure 2. These results give evidence of the value of AES for forecasting end-of-year student performance.

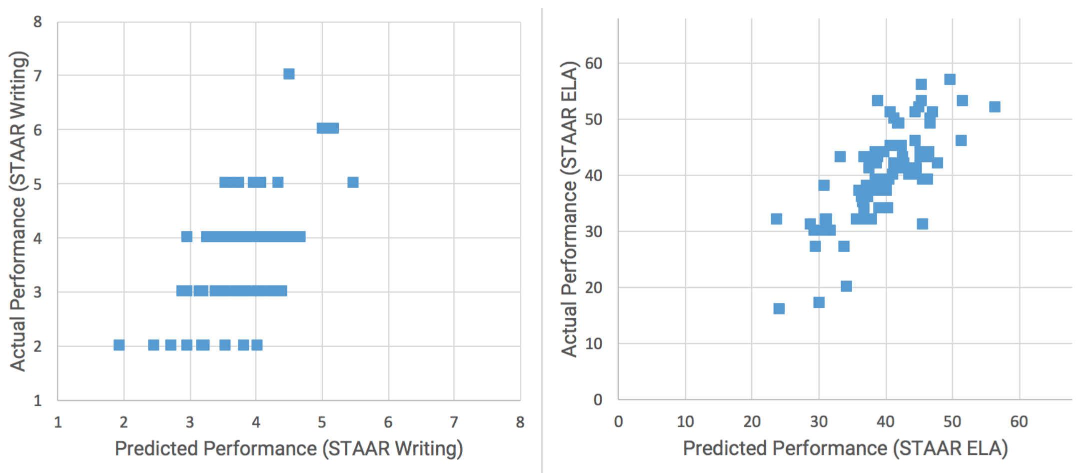


Figure 2: 4-Variable Forecast of scores on STAAR Writing (left) and STAAR ELA (right).

4 IMPROVING SCHOOL OUTCOMES

The next study evaluates a more challenging benchmark: whether use of AES within a standard ELA curriculum improves outcomes over time. To study this, we conducted a multi-site, quasi-experimental study of middle schools in a large, rural school district in Maryland.

4.1 Methods

Teachers in the school district were provided with unlimited access to Revision Assistant during the school year. In fall 2016, trainings were conducted on-site in large groups, and virtually in smaller groups. Staff provided resources to teachers that aligned content in Revision Assistant to school curricula. No specific pacing was mandated by the district. Five schools participated in a treatment condition using Revision Assistant in their curriculum, while two schools in the district did not participate.

Maryland is a consortium member of the Partnership for Assessment of Readiness for College and Careers, or PARCC, which authors end-of-year assessments based on the Common Core State Standards (Maryland Department of Education, 2017). School performance was measured using the PARCC end-of-year assessment for English Language Arts students in 8th-grade, the final year before entrance to high school. Students who exceed or meet expectations, the top two performance categories, are defined as passing.

Table 2: Usage statistics for five participating schools in the 2016-2017 school year.

School	Signal Checks	# Drafts / # Submissions	Mean Increase in Summed Score	Mean Increase in Word Count
1	2,011	14.1	5.5	636
2	3,187	9.9	2.7	788
3	596	5.6	1.9	194
4	6,155	11.3	3.0	413
5	4,733	11.0	2.1	294

Table 3: Change in 8th-grade ELA pass rates in treatment schools.

School	2016	2017	Change
1	35.8	49.1	+13.3%
2	41.1	49.0	+7.9%
3	39.0	45.8	+6.8%
4	30.7	35.5	+4.8%
5	23.7	22.2	-1.5%
Treatment Average (n=5)	34.1	40.3	+6.2%
Non-Treatment Average (n=2)	40.4	38.7	-1.7%
Maryland Average (n=352)	31.8	32.0	+0.2%

4.2 Results

Table 2 describes Revision Assistant usage overall within participating schools. Four schools show high activity while one shows lower levels of activity. In all schools, students composed many drafts prior to submitting their work, represented by the drafts per submission count. The summed score is summed over rubric traits, and so is on a scale of 4 to 16. Mean number of drafts per student submission, mean increase in summed score, and increase in word counts all broadly replicate the finding from Woods et al (2017) of students receiving automated feedback and subsequently improving their essays; however, any one of these measures in isolation is incomplete in capturing student growth or essay quality.

Table 3 presents performance of the five schools in the PARCC exam. Average increase among these schools was 6.2%. By contrast, passing rates declined in both non-treatment schools in the district, by 0.2% and 3.2%. Statewide, average change in pass rates from 2016 to 2017 was an increase of 0.2% (PARCC did not administer an exam prior to 2016, so no further historical data is comparable). The scatter plot in Figure 3 places these schools in the statewide context of all 352 middle schools.

To evaluate the significance of the high rate of improvement in treatment schools, we conducted a permutation test, randomly sampling subsets of five schools from the full population. This lets us evaluate the probability of five arbitrary schools showing similar growth by chance, though it does not account for any potential confounding factors driving both testing success and Revision Assistant usage. These sampled subsets of schools showed mean change greater than 0% in 51% of random samples, amounting to a coin flip. Permutation subsets of schools with mean growth over 6.2% were rarer. Subsets matched or exceeded the performance observed in Revision Assistant schools in 6% of simulations, indicating a 6% chance of these results being observed by chance.

5 DISCUSSION

Combined, these two studies present evidence of the effectiveness of AES in school settings. The first study demonstrates forecasting power of AES to predict student outcomes. The second study demonstrates moderate evidence for improved outcomes and student growth tied to the deployment of an AES product in classrooms during a school year.

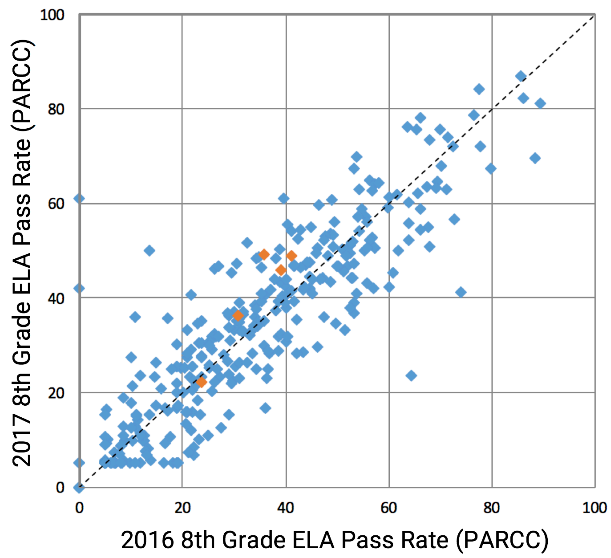


Figure 3: Performance growth of 5 treatment schools (orange) against all other MD schools (blue). The diagonal dashed line represents no year-over-year change.

Based on these results, we recommend two possible paths for schools using AES for forecasting purposes. For locations capable of administering and collecting a full benchmark replication of end-of-year assessments, that process continues to have value. When combined with formative AES activity, the overall predictive accuracy is high for student writing performance and very high for overall ELA performance.

However, the administration of a full benchmark assessment is time-consuming and distracting. For schools without the resources or time for these benchmarks, the results in this work suggest that a single, lightweight AES assessment is a moderately reliable indicator on its own, and adds minimal scoring overhead. In either case, these results are available early in the school year and provide time for targeted intervention. Moreover, the results of the Maryland study suggest a positive impact of AES in year-long classroom use.

5.1 Limitations and future work

Both studies have significant limitations. Neither study was subject to random assignment, relying on volunteers and self-selection. Furthermore, teachers were not subject to a rigorous pacing guide. The impact of AES may therefore be confounded with pre-existing differences, such as teacher readiness for adoption of educational technology, variations in funding of individual schools, or preparedness of building-level instructional coaching staff. Further research will require replication of these results with controlled assignment of students to conditions. The collection of student metadata will remove additional confounds and allow the evaluation of AES systems in light of recent work in fairness of machine learning systems (Leidner & Plachouras, 2017). Furthermore, future research could investigate *individual* student outcomes, an even more granular result not yet studied here.

REFERENCES

Attali, Y. & Burstein, J. (2006). Automated Essay Scoring With E-rater V. 2.0. *Journal of Technology, Learning, and Assessment (JTLA)*, *4*(3).

Crossley, S. A., et al (2013). Using automated indices of cohesion to evaluate an intelligent tutoring system and an automated writing evaluation system. *International Conference on Artificial Intelligence in Education* (pp. 269-278). Springer, Berlin.

Dix, S. (2006). “What did I change and why did I do it?” Young writers' revision practices. *Literacy*, *40*(1), 3-10.

Graham, S., & Harris, K. (2005). *Writing Better: Effective Strategies for Teaching Students with Learning Difficulties*. Brookes Publishing Company. Baltimore, MD.

Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *Journal of Technology, Learning, and Assessment*, *8*(6).

Hayes, John & Flower, Linda. (1980). Identifying the organization of writing processes. Cognitive Processes in Writing (3).

Leidner, J. L., & Plachouras, V. (2017). Ethical by Design: Ethics Best Practices for Natural Language Processing. *European Association for Computational Linguistics 2017*, 8.

Maryland Department of Education. (2017). PARCC for Parents & Teachers. Retrieved from http://www.marylandpublicschools.org/programs/Pages/Testing/PARCC/index.aspx

Page, E. B. (1966). The imminence of... grading essays by computer. *The Phi Delta Kappan*, *47*(5), 238-243.

Riedel, E., Dexter, S. L., Scharber, C., & Doering, A. (2006). Experimental evidence on the effectiveness of automated essay scoring in teacher education cases. *Journal of Educational Computing Research*, *35*(3), 267-287.

Roscoe, R. D., Allen, L. K., Weston, J. L., Crossley, S. A., & McNamara, D. S. (2013). The Writing Pal intelligent tutoring system: Usability testing and development. *Computers and Composition*, *34*, 39-59.

Scharber, C., Dexter, S., & Riedel, E. (2008). Students’ experiences with an automated essay scorer. *The Journal of Technology, Learning and Assessment*, *7*(1).

Scherff, L., & Hahs-Vaughn, D. L. (2008). What we know about English language arts teachers. *English Education*, *40*(3), 174-200.

Shermis, M. D. (2014). State-of-the-art automated essay scoring: Competition, results, and future directions from a United States demonstration. *Assessing Writing*, *20*, 53-76.

Shermis, M. D., & Burstein, J. (Eds.). (2013). *Handbook of automated essay evaluation: Current applications and new directions*. Routledge.

Shermis, M. D., Garvan, C. W., & Diao, Y. (2008). The Impact of Automated Essay Scoring on Writing Outcomes. *National Council of Measurement in Education.* New York, NY.

Texas Education Agency. (2017). STAAR Performance Standards. Retrieved from https://tea.texas.gov/student.assessment/staar/performance-standards/

Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2011). Relating self reports of writing behaviour and online task execution using a temporal model. *Metacognition and Learning*, *6*(3), 229-253.

Williamson, D. M., Xi, X., & Breyer, F. J. (2012). A framework for evaluation and use of automated scoring. *Educational measurement: issues and practice*, *31*(1), 2-13.

Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers & Education*, *100*, 94-109.

Woods, B., Adamson, D., Miel, S., & Mayfield, E. (2017). Formative Essay Feedback Using Predictive Scoring Models. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining* (pp. 2071-2080). ACM.

Adopting Learning Analytics to Enrich Regular Curriculum Review and Enhancement: A Case Study of a University English Programme in Asia

Julia Chen
The Hong Kong Polytechnic University
julia.chen@polyu.edu.hk

Dennis Foung
The Hong Kong Polytechnic University
dennis.foung@polyu.edu.hk

Christine Armatas
The Hong Kong Polytechnic University
christine.armatas@polyu.edu.hk

ABSTRACT: **Practitioner Presentation.** Learning Analytics (LA) has been widely used in predicting students' success, identifying disengaged students, and understanding students' profiles, but not much has been done to incorporate LA into a regular framework. This presentation describes how a University English Programme consisting of three courses adopted LA into their periodic curriculum review framework to supplement the existing quality enhancement measures with a range of learning analytics studies. With the help of LA, a number of policy-level findings were revealed, such as the progression of students between courses, the effectiveness of ability-grouping policy, the learners' pattern of blended learning tasks and the profile of students in various disciplines. These LA-led reviews brought a new perspective to the curriculum, from an isolated / student-feedback-centered model, to a multilevel / evidence-driven model. In the light of LA results, a number of changes have been made, such as the change of blended learning policy and enhancement of blended learning tasks. Evaluation shows that the LA-enriched review is effective in identifying issues that could not have been identified with a traditional curriculum review approach. In view of this, LA should go beyond a could-have innovation to a regular practice in curriculum review.

Keywords: Learning Analytics; Curriculum Review; Academic English; English Curriculum

1 PROJECT OVERVIEW

This project forms part of a curriculum review that has incorporated LA in an innovative manner. On top of a traditional course-based review team, an LA team was formed as part of the curriculum review group. The team mainly followed up course-based review teams' findings (with teachers', students' and external reviewers' perspectives) with an LA approach. They also investigated a range of issues at the programme level with a data-driven methodology. The use of LA in this project was a success as the evidence-driven approach has generated results that corroborate with the findings from the course-based review team, and identified issues that could not have been found from periodic course-based reviews. The suggestions made were well-grounded in data, and the LA approach led to clear and concrete suggestions for improving the course.

2 BACKGROUND

Hong Kong underwent a round of major curriculum reform in the last few years and the current project is situated in this era of change. The Hong Kong SAR government introduced a 4-year university system in 2012 to replace the old 3-year system. Subsequently, a number of changes were made to the senior secondary school curriculum and the university undergraduate curriculum. The university in question decided to introduce a revamped core English curriculum. All students have to take two English courses to fulfill the "Language and Communication Requirement" as part of the university curriculum. The courses they take depend on their English scores in their secondary school exit exam. The study pathways are illustrated in Table 1. Since the number of students taking these three courses is huge, there are regular quality assurance (QA) measures to ensure the validity and the effectiveness of the course. In 2016, the first batch of students under the new system graduated from their studies and this motivated the host department of these courses to review them in a comprehensive manner.

Table 1 – The progression pattern for different proficiency levels

HKDSE English Language Attainment	***Subject 1***	***Subject 2***
Level 5 or above (IELTS 6.81+)	Advanced English for University Studies (A-EUS)	Advanced Elective Subjects
Level 4 (IELTS 6.31–6.51)	English for University Studies (EUS)	Advanced English for University Studies (A-EUS)
Level 3 (IELTS 5.48 – 5.68)	Practical English for University Studies	English for University Studies (EUS)

Traditional curriculum review is a stakeholder-centered model and LA can bring synergy to this. Traditional review focuses on the perspective presented by different stakeholders via student interviews, teacher interviews and expert review of teaching materials. These perspectives provide useful suggestions for course teams to enhance the courses but the very nature of interviews can lead to subjective results, depending on who was being interviewed, when the interviews were conducted and what and how questions were asked in interviews; and these are in fact common issues in any qualitative research. To introduce a more objective and evidence-based perspective to curriculum review, the current review project adopted LA to triangulate the qualitative findings.

It is important, as an abstract, to note that the LA team focused on the relevant literature when analyzing different aspects of the curriculum. For example, literature on writing progression (such as Archibald, 2001; Storch, 2009) was consulted when the LA team investigated the extent of student improvement.

3 IMPLEMENTATION

This project started with the establishment of different course-based review teams, consisting of course leaders, course teachers and external reviewers. Other than the traditional composition of teams, a learning analytics team was formed as one of the teams, with course teachers, IT specialist, and educational technologists. The LA team actively approached various review teams and followed up various results they found from the review process. The LA team then continued to look for evidence-driven details to triangulate the findings and made suggestions for improvement, such as the deadlines for blended learning task and the usefulness of blended learning activities. Other than course-level matters, the LA team also investigated a number of major matters at the curriculum-level, such as progression of students, and effectiveness of ability-grouping practices. The LA team provided a range of suggestions for the regular review framework for further enhancement.

3.1 Programme LA: Progression of Students between Courses

Student progression was one of the major areas that the review group wanted to investigate. Since all students need to take two courses to fulfill the university requirement, the LA team wanted to explore how students progressed from the first English course to the second English course. Past literature on writing progression shows that previous studies did not have large samples of students because it was hard to recruit participants to trace their development. Retrieving data from the university's learning management system, the IT colleagues in the LA team merged the results of the two courses and compiled a full list of students for each study pathway. Since the second assignment of the two courses is comparable with the same set of marking descriptors, the LA team compared the performance of students in two courses. Simple descriptive statistics and Paired-Sample T-tests were performed. The results illustrated that students show better improvement in referencing skills than proficiency-based elements. Such findings allow course designers to plan the learning focus of each pathway.

3.2 Programme LA: Ability-grouping Practices

Streaming, also known as ability-grouping or ability-tracking is not a common practice in higher education but this university has decided to adopt streaming as the English proficiency of its students ranges from lower Level 3 to upper Level 5 at entry. (See Table 1 for IELTS equivalents). As seen in Table 1 above, some students may enroll in one course as their first English subject, while some other students may take it as their second English subject. Previous studies on ability-grouping tend to focus on secondary or primary education, and not much has been done with tertiary students. With the help of IT colleagues, the LA team compared the results of different groups of students taking the same course. Simple data visualization techniques and Independent Sample T-tests were used. Results suggest that the two groups of students did not differ in the extent of improvement in language skills, but did differ in skill-based elements. Only conducting LA could have provided insights into this aspect.

3.3 Course LA: Deadlines of Blended Learning Tasks

The LA-enriched review has helped course leaders to set a better deadline for blended learning tasks. In one of the courses, students are required to complete one blended learning task each week (i.e. 13 tasks for the whole course) and the deadline for most tasks was at the end of the course. These tasks may include watching an online video and completing multiple-choice questions to reinforce what students have learnt each week. Through the traditional quality assurance mechanism, it was found that students tend to complete all tasks at the end of the course and such last-minute behaviour did not contribute much to their learning. By conducting network analysis and visualizing the completion patterns, LA confirmed the findings in the traditional review. Decision tree analysis showed that students who complete their tasks earlier achieve more in this course, i.e. they have higher grades. The findings provide empirical justification for course leaders to change the deadlines of online learning tasks to three staggered deadlines for completion of online learning tasks so as to discourage last-minute completion.

3.4 Course LA: Relationship between Blended Tasks and Learning Outcomes

Another contribution of LA to the course review is in its ability to show how various blended learning tasks can predict learning outcomes. The course-based review team believe that the blended learning tasks are meaningful to students' learning after analyzing the content of the questions. However, the use of LA can show which learning task actually predicts the learning outcomes of students. By conducting several rounds of Multiple Regression analyses, results of several online learning tasks stand out as being able to predict students' outcome. These results once again provide course designers and course leaders with another perspective and help to facilitate the process of adapting / removing certain online learning tasks.

4 IMPACT AND EVALUATION

This curriculum review project has several important implications for this university English programme. First, with the results of the progression and streaming studies, it was decided to include more language-focused blended learning tasks so that better language improvement could be made. Second, online learning tasks now include three staggered deadlines, instead of one end-of-term deadline. Two small scale follow-up studies were conducted to evaluate the effectiveness of these changes and positive results were found. The insights generated from our LA analyses warrant the suggestion of incorporating LA into regular curriculum review to form a multi-layered, evidence-based evaluation framework.

REFERENCES

Archibald, A. (2001). Targeting L2 writing proficiencies: Instruction and areas of change in students' writing over time. *International Journal of English Studies*, *1*(2), 153-74.

Storch, N. (2009). The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing*, 18(2), 103-18.

Hai Linh Truong, flickr.com

Practitioner Posters

University Leadership and Support in Learning Analytics

Jess TAN Wei Chin
Singapore University of Social Sciences
jesstanwc@suss.edu.sg

Gabriel GERVAIS
Singapore University of Social Sciences
gabrielgervais@suss.edu.sg

KOH Hian Chye
Singapore University of Social Sciences
hckoh@suss.edu.sg

ABSTRACT: This paper shares how senior management in the Singapore University of Social Sciences (SUSS) embedded learning analytics into its strategy map, and teaching and learning. For several years, learning analytics projects were undertaken via funded research projects. This changed in August 2016 when a dedicated unit was set up to promote institutional analytics as a strategic resource to support data-driven and evidence-based decision making. With this, learning analytics became more important to the University's core business of higher education. Consequently, the 2016 initiative led to the setting up of a data warehouse to facilitate analytics, the incorporation of analytics into the University's strategic plans, the provision of analytics training for faculty and staff, and the implementation of learning analytics projects. While SUSS is now seeing results from the 2016 initiative, the progress made to date is not without difficulties. These include getting buy-in from stakeholders, ensuring data integrity and building up the University's analytics capability. Senior management leadership and support is the most critical factor for success. It is hoped this paper can benefit other universities intending to embark on the same journey.

Keywords: Learning analytics, institutional analytics and research, university leadership in analytics, the learning analytics journey, difficulties in implementing learning analytics

1 BACKGROUND

In 2011, the first LAK Conference defined learning analytics as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environment in which it occurs" (LAK, 2011). Grama (2015) reported that learning analytics was already and increasingly being used by universities in institutional business and student engagement and learning, and these had led to improved experiences for students, faculty, staff and others; as well as improvements in programmes and courses. (See also Grajek (2014) for similar conclusions).

2 EMBEDDING LEARNING ANALYTICS

For several years, learning analytics projects were among many other university funded-projects in the Singapore University of Social Sciences (SUSS). In 2016 however, to address its need for data-driven and evidence-based decision-making and planning, SUSS senior management included learning analytics as one of the key dimensions of the University strategic thrusts.

3 IMPLEMENTING LEARNING ANALYTICS

As a result, SUSS established the Institutional Research Analytics Unit (IRAU) in August 2016 that has since been leading three main initiatives: setting up the data analytics infrastructure and resources, developing the data analytics capabilities through staff training and undertaking strategic, tactical and operational institutional analytics projects. The 2016 analytics initiative can be represented by Figure 1 below.

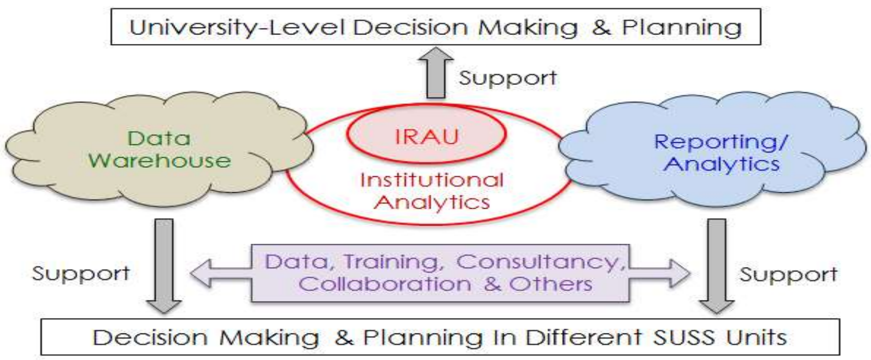


Figure 1: Analytics Framework in the Singapore University of Social Sciences

4 DIFFICULTIES, FUTURE DIRECTIONS AND CONCLUSION

While the implementation of institutional and learning analytics met with a number of challenges (arising, among others, from data integrity, accountability and responsibility issues as well as significant diversity of staff profile, job scopes and motivation levels), it remains critical that more efficient and effective decisions are made about teaching and learning support leading to better learning outcomes and experience for the students at the university. The progress made to date to embed learning analytics at SUSS could not have been realised without the support and leadership of the University's senior management, undoubtedly the most critical success factor of this initiative.

REFERENCES

Grajek, S. (2014). Top ten issues, 2014: Be the change you see. *Educase Review Online*. https://net.educause.edu/ir/library/pdf/ERM1421.pdf

Grama, J. L. (2015). *Developing strategies for implementing analytics. EDUCAUSE Review Online*. http://www.educause.edu/ero/article/developing-strategies-implementing-analytics

LAK (Learning Analytics and Knowledge). (2011). https://tekri.athabascau.ca/*analytics/ call-papers*

Towards a Semi-open Data Pipeline for Practice and Research

Adam R Cooper
Tribal
adam.cooper@tribalgroup.com

ABSTRACT: This **poster** outlines progress made towards a standards-based data pipeline for learning analytics which accounts for the needs of a large scale multi-tenant operational learning analytics solution. It takes a further step by identifying opportunities which the arising data-centric architecture may afford, specifically opportunities from making the data pipeline open and accessible to authorized researchers in the institution for whom the solution is operated. Potential opportunities include, for example: transparency and accountability, reduced effort for researchers on mundane data processing, and an easier pathway for research findings to propagate to at-scale operational learning analytics. The poster presenter will solicit clarified and augmented opportunities from LAK participants and seek votes on their relative importance and desirability, with a view to implementing them.

Keywords: standards, transparency, accountability, scalability, data, infrastructure

1 BACKGROUND AND CONTEXT

Among the technical challenges for the deployment of learning analytics at scale, and for a sustained period of time, is the preparation of data. This must require minimal user intervention to be cost-effective and quality-assured. It should also be: scalable, resilient, robust, maintainable, and extensible. The approach outlined in section 2 has been put into practice through a series of projects involving mid-size UK institutions participating in the Jisc Effective Learning Analytics programme[1].

Ongoing discourse in the learning analytics community around two themes has suggested further opportunities. These themes are: the transparency and accountability of learning analytics, especially of commercial products; the desire for more impact of learning analytics research on at-scale operational systems. We suggest that s*emi-open data* - by which we mean the opening-up of a commercial data pipeline to authorized customer staff – can help. Various kinds of data sharing have been surveyed in the LACE Project publication of "Data Sharing Requirements and Roadmap" (Cooper & Hoel, 2015) but we are not aware of existing initiatives with the same approach described here. The University of Michigan Learning Analytics Data Architecture[2] opens-up its data warehouse to that institution's researchers, but our approach would provide access to more refined data.

2 ARCHITECTURE OF THE STANDARDS-BASED DATA PIPELINE

Our approach has been to establish a data pipeline, comprising numerous atomic data processing stages with a well-defined set of dependencies.

[1] https://www.jisc.ac.uk/rd/projects/effective-learning-analytics
[2] https://enrollment.umich.edu/data-research/learning-analytics-data-architecture-larc

A simplified view of part of the data pipeline is shown in Figure 1[3]; other parts deal with "activity data". The previously-outlined design aims are dealt with as follows:

- cost-effective – local data structures are mapped to the Jisc UDD[4] administrative data standard, enabling a consistent treatment in the down-stream pipeline for all clients;
- quality-assured – the processing of dependencies is fully automated, with input/output logging at each step;
- scalable – atomic tasks can be scheduled to spread load over time;
- resilient and robust – failure is at atom-level and each step validates its input;
- maintainable – due to steps with clear functional bounds;
- extensible – low impact from atom addition/extension.

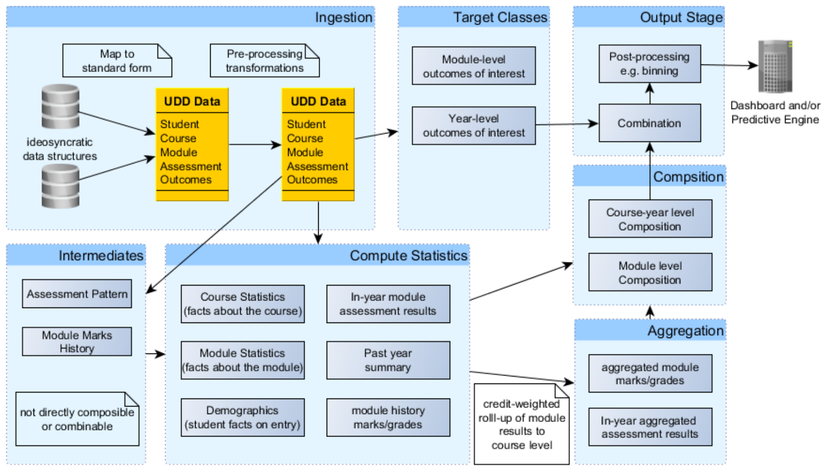


Figure 1: Simplified view of the UDD part of the data pipeline

3 TOWARDS A SEMI-OPEN PRACTICE + RESEARCH PARADIGM

Opening-up the data at every stage of processing, right up to the door of the predictive engine and dashboard, has benefits, subject to a trusted privacy and security model, which this poster seeks to engage LAK participants in discussion over:

- Transparency and accountability of a commercially-operated platform becomes achievable. Deep transparency is not tractable via user-centered dashboard tools.
- Reduced researcher effort dealing with routine data preparation.
- A shared data dictionary between researchers, teams, and potentially in inter-institutional efforts, which would support replication and script-sharing.
- There is an uncomplicated pathway for research findings into operational systems.
- It would stimulate greater research-practitioner collaboration within the institution.

REFERENCES

Cooper, A., Hoel, T. (2015, June). *Learning Analytics Community Exchange D7.2 – Data Sharing Requirements and Roadmap.* Retrieved from http://www.laceproject.eu/deliverables/d7-2-data-sharing-roadmap/

[3] The diagram uses short-hand "course" and "module" for units of study typical in higher education; a course comprises many modules and runs over one or more years to a final award.
[4] https://github.com/jiscdev/analytics-udd/

Use of the Virtual Learning Environment (VLE) and lecture recording does not correlate with student performance in first year Biology modules

Author(s): Please Leave This Section Blank for Review
Institution
Email

ABSTRACT The aim of this study was to determine if engagement with the VLE, in particular lecture recordings, improved students exam performance. Online material including widespread provision of lecture recordings is increasingly requested by students. However there is little appreciation of how this material is used by students and if it improves performance. In two first year Biology modules, analytics relating to use of the Virtual Learning Environment (VLE; Canvas) usage was compared to module marks. There was no strong correlation between either engagement with the VLE or viewing of lecture recordings. Interestingly there were no significant differences in the patterns of usage between high performing students (Quartile 1) and those at the bottom of the scale (Quartile 4). However Q1 students engaged earlier and more consistently with lecture recording in the run up to exams. The results suggest a complex relationship between VLE usage and students performance. While lecture recordings are reported as useful for some students groups (e.g. those with disabilities or language issues) further work is needed to understand how students use the VLE and how academic staff can best support the student experience to facilitate good performance.

Keywords: Lecture recording, VLE, student performance,

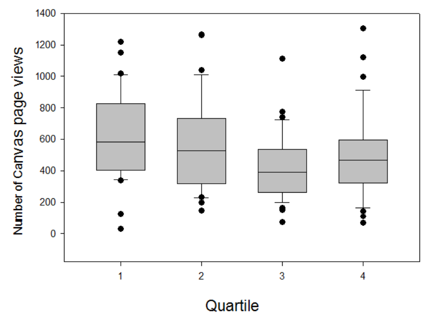
1. The increasing use of the VLE and lecture recoding

The provision of online resources including lecture recording implies it has value to learning. In the absence of other information students may assume that use of such resources will lead to an increase in performance. However the evidence is equivocal. A study on large UG classes in Irvine, USA showed no strong correlation between student use of lecture recording and overall performance (Williams *et al.*, 2016). Nursing students with access to lecture recording performed worse than cohorts with only conventional lectures (Johnson, 2013).

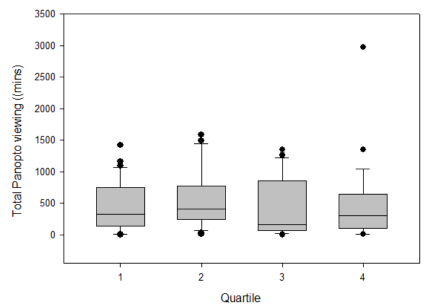
To assess the value of the increased provision of online resources we gathered data on the use of the VLE, lecture recording and student performance for first year UG students on two introductory Biology modules. The VLE contained a wealth of material to support learning, including lecture recordings, POD casts, web links, multiple choice quizzes and exam marking exercises. The final exam was composed of questions of similar format to those in the online formative exercises so a correlation between usage of the VLE and performance might be expected. Spreadsheets were downloaded from the Canvas VLE as 'number of page views' while lecture recording data was obtained as number of times recordings were accessed or total time they were viewed. The student cohort was divided into quartiles based on the final overall module marks and the performance of each quartile compared.

2. COMPARING VLE USAGE BETWEEN DIFFERENT PERFORMING STUDENTS

1a)



1b)



1c)

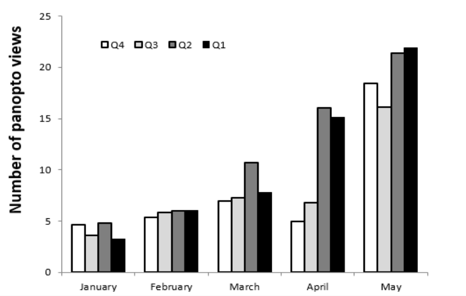


Figure 1: Comparison of engagement with the VLE (1a) lecture recording, (1b) module performance in differently performing quartiles in a core first year biology module. 1c) Number of lecture recording views each month by Q1-Q4 performing students. Kruskal-Walllace showed a significant effect of quartile (P=0.005) with the lower performing quartiles of 1 and 2 viewing lecture recordings less than Q3 and Q4 in the April period. The teaching was delivered January - March, the exam was in June.

3. DISCUSSION

This pilot study demonstrated no strong effect of the use of the VLE or lecture recordings on student performance, although Q1 and Q4 students sued it differently for revision. Similar conclusions were also found comparing first semester (naïve) students with those in the second semester and in subsequent years iterations of the same modules (data not shown in abstract but will be presented in the poster). Further work will explore the ways which in students are the resources differently *e.g*. Luttenberger *et al.,* (2017) found that high performing students are more selective in how they use this material. High performing students viewed lecture capture less and had a more discerning pattern of engagement indicating the potential diversity of usage amongst students who learn in different ways (Owston *et al.,* 2011). As a result of this study these analytics are now used in sessions introducing students to VLE and lecture capture to help them to reflect on how they are using these resources.

4.REFERENCES

Johnston, A.N.B.; Massa, H. and Burne, T.H. J. (2013) Digital lecture recording: A cautionary tale. *Nurse Education in Practice* 13 40-47

Luttenberger, S., Macher, D., Maidl, V., Rominger, C., Aydin, N. and Paechter, M. (2017). Different patterns of university students' integration of lecture podcasts, learning materials, and lecture attendance in a psychology course *Educ Inf Technol*. https://doi.org/10.1007/s10639-017-9592-3

Owston, R., Lupshenyuk, D. and Wideman, H. (2011) Lecture capture in large undergraduate classes: Student perceptions and academic performance. Internet and Higher Education 14 262-268

Williams, A.E., Aguilar-Roca, N.M. and O'Dowd D.K (2016). Lecture capture podcasts: differential student use and performance in a large introductory course. *Education Tech Research Development, 64*, 1-12. http://dx.doi.org/10.1080/1359432x.2013.877892

Using Learning Analytics to Evaluate Course Design and Student Behavior in an Online Foundations of Wine Science Course

Kerry Wilkinson, Imogen McNamara, Karina Riggs, David Wilson
The University of Adelaide
karina.riggs@adelaide.edu.au

ABSTRACT: This study evaluated the approach to course design employed to transition a course from face-to-face to online delivery. Traditional lectures were initially replaced by audio recordings, and subsequently to purpose built eLearning video lectures with online resources aimed to engage, stimulate, support and enhance student learning. Descriptive learning analytics, together with student grades and student based learning and teaching evaluations, indicated the use of audio recordings had a negative impact on the perception of course organization, learning strategies, and course quality. The restructured course with purpose built online video lectures was considered by students to be superior to audio recordings and equivalent to face to face delivery. Despite student perceptions there was no difference in student learning outcomes.

Keywords: learning analytics, online learning, learning outcomes, student engagement.

1 INTRODUCTION

Learning analytics (LA) refers to the measurement, collection, analytics and reporting of data about the progress of learners and the context in which learning takes place (Siemens, 2012). The Learning Analytics in Higher Education Review states, LA could make significant contributions to: quality assurance and improvement; retention rates; assessing and acting upon differential outcomes among the student population; and the introduction of adaptive learning (Scatler, et al. 2016). Adaptive learning systems are emerging to help students develop skills and knowledge in a self-paced way. Clickstream data can be used by educators to evaluate how content is used and how effectively it supports student learning. LA has the potential to transform the way learning outcomes and impact are measured and to inform the development of modern approaches to achieving excellence in teaching and learning. This study used LA to evaluate student engagement, learning behavior and outcomes in a Wine Business course as it transitioned from face-to- face to online.

2 COURSE STRUCTURE AND STUDENT DEMOGRAPHICS

Foundations of Wine Science (FWS) comprises both theoretical and practical components of a Wine Business program. Prior to 2016, the theoretical component of FWS was taught via traditional face-to-face lectures, in 2016, the program transitioned to simple online delivery of lectures as audio recordings (recorded lecture slides with audio). As, a consequence of reduced student evaluations of learning and teaching in 2016 the course content for 2017 was redesigned to include purpose built eLearning video lectures with additional online resources including study guides, tutorials, discussion boards and learning interactives. The majority of FWS students were international (≥85%), predominantly from China; a significant proportion of whom had limited English literacy.

Demographics impact the way in which students interact, navigate and engage with online materials, providing insight into patterns of behavior for learners.

3 LEARNING ANALYTICS AND STUDENT BEHAVIOUR

Transitioning this course from face-to-face to audio lectures and purpose-built eLearning video lectures did not impact students' final grades; (72.2, 71.1, 71.7, percent respectively). Data from student evaluations of learning and teaching identified a positive shift in engagement when course content moved from audio to purpose built eLearning video lectures. The LA data revealed 55% of students accessed the video lectures, while 100% of students accessed lecture notes in PDF format. The lower percentage of video views may reflect the high percentage of international students in the course and their learning preference for text based materials rather than non-closed captioned video. These data highlight the value of providing learning resources in a variety of formats to accommodate student learning preferences. The provision of closed captions/transcripts may expand the value of video lectures. When moving to the eLearning video lecture format, discussion boards were introduced to enable students to engage with the lecturer. A total of 24 posts were made to the discussion board with 82% of students viewing these within 2 hours of posting. Although less than 10% of students were responsible for the majority of posts, 76% of students engaged with the discussion board in a timely manner. Interestingly a large proportion of students re-engaged with posts several times perhaps indicating they were reflecting and clarifying their understanding as the course went on. The fact that posts were not anonymous may also have influenced the extent of active student participation in discussion boards (Roberts & Rajah-Kanagasabai, 2014). Moving to a new LMS afforded improvement in course structure reducing the number of clicks required for course navigation (from 4 to 1-2). The data indicated course announcements, assignments, discussions and marks were the most highly viewed pages. In contrast, feedback was viewed by less than 25% of students, indicating there may be value in encouraging students to engage with feedback to provide the necessary remediation to meet their learning outcomes.

4 CONCLUSION

The LA indicate that students undertaking FWS have a range of learning preferences, with the majority of students preferring text-based resources. This was not surprising, given the high proportion of foreign and ESL students (>85%). Improvements in student evaluations from 2016 - 2017 may be attributed to: improved quality of video lectures (engaging images, captions, high production value); simplified structure; and increased diversity in learning resources.

REFERENCES

Roberts, L. D., & Rajah-Kanagasabai, C. J. (2014). "I'd be so much more comfortable posting anonymously": Identified versus anonymous participation in student discussion boards. *Australasian Journal of Educational Technology*, *29*(5), 612-625.

Scatler, H., Mullan, J., & Peasgood, A. (2016). Learning Analytics in Higher Education: A review of UK and International practice. JISC.

Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. Paper presented at the 2nd International Conference on Learning Analytics and Knowledge (LAK12), Vancouver, Canada.

Early prediction dropout through assignment delivery in adult e-learning course

ABSTRACT: In this paper, a dropout prediction analytics strategy for adult e-learning courses based on assignment and participation student data, is proposed. Despite the fact that a great number of students enrolled in the course, a 30% of student dropout were reported. Since we observed from data that student dropout starts at the middle stage of the course, the aim of this work is to determine in which number of assignment can student dropout be predicted to generate retention alerts.

Keywords: E-learning, Tertiary education, Adult e-learning, Dropout prediction.

1 INTRODUCTION

In 2010, at the National University of San Luis, a Tertiary degree in University Institutions administration and management was launched under face-to-face modality. In the year 2015, the third cohort was launched with a virtual education proposal. The target population is non-teaching staff of the university.
In this paper, we present the data analysis of Reading Comprehension course corresponding to the first year of the career. This analysis represents a first approximation to the use of learning analytics on a course in our university. In this course 216 students participated.

2 COURSE CHARACTERISTIC

The course used Moodle platform as a space for study and participation.
Students presented heterogeneous characteristics based on age, educational level, digital competences and functions. With respect to educational level, some students were in the process of finishing High school and others were professionals in hierarchical functions.
With relation to digital competition, some did not have an Internet or computer connection in their home, others did not have information and communication basic processing skills. In opposite situation, we had students with advanced digital skills. Another determining characteristic for dropout was the time asign to study.

3 DATA PREPARATION

The activities were carried out in different environments, so a follow-up sheet (XSLS format) was designed for student participation.
For this work the data was obtained from group activities in Google Docs (Practical Work), activities in Socrative, Google Forms (Test), QuestBase (self-assessments).

Participation is recorded in all activities. In addition, in Practical Works (PW) from 2 to 7 the correction provided by the tutor is recorded. The following figure shows the type of activity.

Tools	Activities	Participation	Autotesting	Assigment
Google Docs	PW2	✓		✓
	PW3	✓		✓
	PW4	✓		✓
	PW5	✓		✓
	PW6	✓		✓
	PW7	✓		✓
Socrative	PW1	✓	✓	
Google Forms	Test of Study Habits	✓		
	Learning Styles Test	✓		
QuestBase	self-assessments1	✓	✓	
	self-assessments2	✓	✓	

Figure 1: Type of activities and evaluation system

In order to obtain the minable view, the follow-up table and notes obtained in the (PW) were used. In addition, the following codification of participation in activities was defined:

Table 1: Coding system

Final condition of student	Non-evaluative activities	Grade on evaluative activities
Approved (A)	Do not participate (0)	Very Good (MB)
Not Approved (NA)	Participate (1)	Good (B)
Dropout (AB)		Regular (R)
		Wrong (M)
		Not done (NP)

The minable view consists of 216 records, divided into: 140 Approved (A), 12 Not Approved (NA) and 64 dropout (AB). We will have 3 classes (A, NA, AB) and a total of 19 attributes. Although the NA class has few records, our interest is focused on the AB class, to predict dropout in advance.

The following table shows the structure of the minable view.

Table 2: minable view

Género	1E	P-PW2	N-PW2	P-PW3	N-PW3	P-PW4	N-PW4	P-PW5	N-PW5	P-PW6	N-PW6	P-PW7	N-PW7	P-AE1	N-AE1	P-AE2	N-AE2	TA	Condición
F	1	1	B	1	B	1	B	1	MB	1	B	1	B	1	B	1	MB	1	A
M	1	1	B	1	B	1	B	1	B	1	MB	1	MB	1	R	1	MB	1	A
M	1	1	MB	1	B	1	B	1	MB	1	B	1	R	1	R	0	NP	1	A
F	1	1	MB	1	B	1	B	1	MB	1	B	1	B	0	NP	0	NP	1	A
M	1	1	B	1	B	1	MB	1	MB	0	MB	1	MB	1	B	0	NP	1	A

The records distribution in the Percentage split was 78% for training data (169 records) and 22% for testing (47 records).

The sets were as follows:

Table 3: Training and test sets

	A	NA	AB
Training	112	9	48
Test	28	3	16

Data were divided into 5 training sets based on PW. The objective was to evaluate in which PW instance the dropout can be predicted.

The following table shows how the training sets were organized. PW 6 and 7 are not included because are assignments deliver near the end of the course. Training 1 consists of all PWs, tests and self-assessments.

Table 4: training sets by PW

Sets	Training 1 (T1)	Training 2 (T2)	Training 3 (T3)	Training 4 (T4)	Training 5 (T5)
PW	PW2- PW7	PW2	PW2-PW3	PW2- PW4	PW2- PW5

4 RESULTS

Weka version 3.8.1 is used. A first analysis was performed using the classification algorithm J4.8. The following table shows the accuracy and TP rate values for the AB class of each training.

Table 5: Accuracy & Tp rate

	T1	T2	T3	T4	T 5
Accuracy	97,87%	93,61%	89,36%	93,61%	93,61%
TP rate AB	1	1	87%	1	1

It can be seen that the value for T2, T4 and T5 are the same. As the objective of the analysis is to determine which is the first PW that allows to predict AB, by the values of accuracy one could conclude that the T2 would fulfill this objective. However, in this instance, most students comply with the delivery and approval of PW2, which invalidates it as a relevant PW for our analysis. The decision tree for T1 has PW7 as root, the last PW, then analyzes the PW4 and PW5 that are developed in the middle stage of the course. It does not take into account PW2.

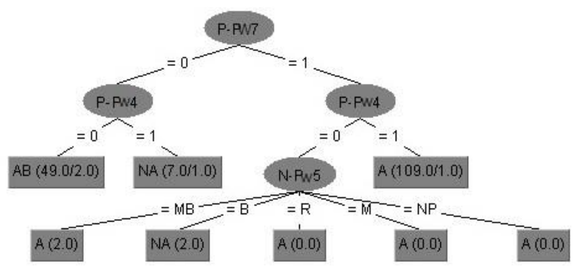


Figure 2: Training 1 Decision tree

Subsequently, a second analysis was performed applying cross-validation with 3 folds.

Table 6: Accuracy & Tp rate

	T1	T2	T3	T 4	T 5
Accuracy	95,18%	90,36%	90,36%	89,15%	89,15%
TP rate AB	97%	86%	88%	88%	88%

Analyzing TP rate provided from T2 to T5, it can be observed that from T3 the values improve with respect to T2 and it remains stable. With this, it can be established that in order to predict dropout, the analysis must be performed with data obtained since PW3.

5 CONCLUSION

In this paper the aim is to determine in which number of PW students can be predicted to dropout the course.

The data analyzed were obtained after the course ended. In the analysis with Weka, it can be observed that between Practical Work Nº3 and 4, dropout can be predicted and is a relevant indicator to generate retention alerts.
As a results, in next course cohort data analysis can be applied during the course development and perform early interventions to avoid dropout.

REFERENCES

López Molina, J. M., Herrero, J. G. (2006). Técnicas de Análisis de Datos. Aplicaciones prácticas utilizando Microsoft Excel y Weka.

Tan P., Steinach M., Kumar V. (2006). Introduction to Data Mining.

Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). Introducción a la Minería de Datos. Madrid: Pearson Prentice Hall.

Tecnicatura en Administración y Gestión de Instituciones Universitarias (Tagiu). Universidad nacional de San Luis. http://planesestudio.unsl.edu.ar/index.php?action=car&fac=6&car=79&plan=19/10&version=6

Socrative. https://www.socrative.com/

QuestBase. http://www.questbase.com/

Advancing the Delicate Issue of Ethics and Privacy for Learning analytics

Tore Hoel and Weiqin Chen
OsloMet – Oslo Metropolitan University
{Tore.Hoel,Weiqin.Chen}@oslomet.no

ABSTRACT: Stakeholders need systematic procedures to walk through and address the different issues of ethics and privacy for learning analytics, moving beyond the DELICATE checklist first presented at the LAK16 conference. This paper presents a process- and role-based framework for managing a structured discourse on these issues. The Toolkit supports groups in setting up a process identifying perspectives and concerns, exploring issues through a systematic walk-through of the learning analytics processes cycle specified in the ISO/IEC 20748 standard.

Keywords: Learning analytics, ethics, privacy, data protection, data management

1 INTRODUCTION

The DELICATE checklist, presented by Drachsler and Greller at LAK16, has been acknowledged by the learning analytics (LA) community as a useful means to "to support educational organisations in becoming trusted Learning Analytics users" (Drachsler & Greller, 2016). The DELICATE checklist is now outdated. When designing a new "reflection aid" (ibid.), it is our view that one should leave the checklist approach as this does not support communication on complex issues (Catchpole, 2015). This paper is the first step to develop a scaffold for systematic discussion on ethics and privacy for LA.

2 INITIAL DESIGN OF THE EP4LA TOOLKIT

The aim of this design is to develop a set of tools that allow different stakeholders to explore ethics and privacy issues related to LA and to contribute to solutions. We see this as part of a continuous quality assurance process for LA, in which the principle of data protection by design and by default (European Commission, 2016) is followed, and examining of ethics and privacy issues are carried out on a regular basis. The ethics and privacy for LA (EP4LA) toolkit will be useful in different settings, from occasional workshops organised by the LA research community to more structured institutional processes implementing new LA solutions.

The following steps and tools are suggested for further testing:

1. Who do you represent? What are your aims? The first step is to clarify stakeholder perspective and interests. Identifying loyalties and aims are essential steps in both the ROMA approach (www.roma.odi.org) and in the Potter Box approach (Backus & Ferraris, 2004).

2. Agreeing upon main concerns about ethics and privacy related to LA? As privacy is seen as a major stumbling block to large-scale implementation of LA (Hoel, & Chen, 2016) all kinds of issues may be of concern to stakeholders. Therefore, the point of this step is to agree on an issue to start the exploration. It does not need to be the most crucial threat to privacy, what is sought is more the concern that will kick off the discussion.

3. Brainstorm and map issues related to perspective and concern against the LA process model. In the third step, the discussants will use a template (step III, Figure 1), one per concern.

From the chosen stakeholder perspective, the concern under study is examined using six LA processes as prompts. The idea is that the group brainstorm issues on post-it notes and move them around on the template printout to connect to the different LA processes.

4. Collating concerns and barriers into a first problem space description. When brainstorming issues is reaching saturation the next step is to move the discussion forward towards solutions or policy actions. Again, it is important to decide upon unit of analysis to keep focus, e.g., institution, regional or national level, or a specific stakeholder group. The output of this last step IV (Figure 1) will be a collection of concerns or issues, and the barriers that must be overcome to solve these issues. Ideas for solutions may also be touched upon in the brainstorming that took place in the previous step.

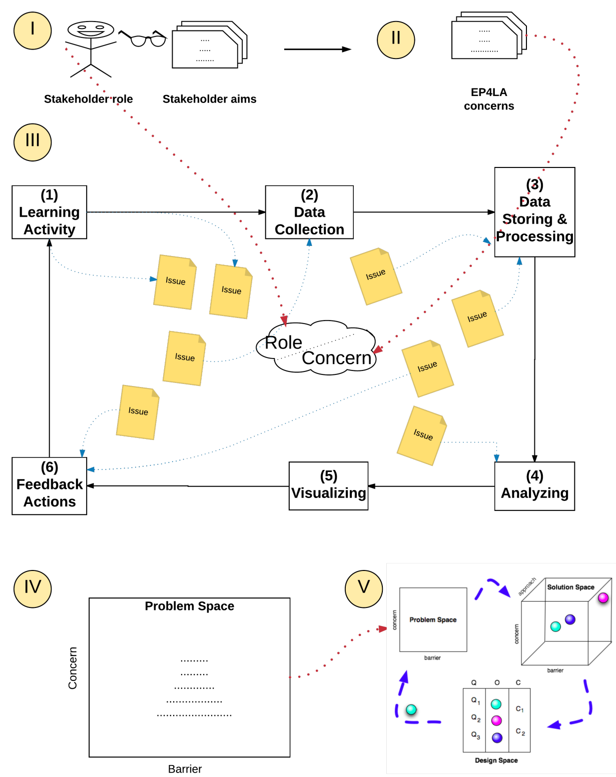


Figure 1: EP4LA toolkit, version 1.0

REFERENCES

Backus, N., & Ferraris, C. (2004). Theory Meets Practice: Using the Potter Box To Teach Business Communication Ethics. In Proceedings of the 2004 Association for Business Communication Annual Convention.

Catchpole, K., & Russ, S. (2015). The problem with checklists. BMJ Quality & Safety, 24(9), 545–549. http://doi.org/10.1136/bmjqs-2015-004431

Drachsler, H., & Greller, W. (2016). Privacy and analytics – it's a DELICATE issue. Presented at the Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16. http://doi.org/10.1145/2883851.2883893

European Commission. (2016). REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

Hoel, T. & Chen, W. (2016). Data Sharing for Learning Analytics – designing conceptual artefacts and processes to foster interoperability. In Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education

Discourse Engagement during Online Test Preparation: A Social Network Analysis Approach

Maria Ofelia Z. San Pedro
ACT, Inc.
sweet.san.pedro@act.org

ABSTRACT: This paper examines student participation in discussion forums within an online test preparation environment as part of ongoing research that looks into online learning activities and assessments that students across the SES spectrum partake in to prepare for high-stakes exams. Social network analysis is employed to evaluate students' engagement in discourse in relation to their eventual exam performance. Preliminary findings presented here provide insights on the potential role of discourse in knowledge-building during online test preparation (especially for underserved learners) and inform future work that can improve the design of the environment.

Keywords: Discussion Forum. Test Preparation. Social Network Analysis. Underserved Learner.

1 INTRODUCTION

Students prepare for exams in a variety of ways – through workbooks, practice tests, classroom curricula, coaching or online tutoring. Research on test preparation suggests that creating more practice opportunities, and instruction-like activities aimed at developing knowledge/skills can have a greater effect on exam performance; and digital environments for learning/assessment have catered to this in recent years. Moreover, computer-supported learning has shown that learner discourse contributes to their knowledge construction. Wang et al. (2015) found that active/constructive online discussion behaviors (in MOOC discussion forums) are predictive of learning gains. Another study showed that active participants in discussion forums have also used other platform activities (videos, quizzes, etc) to a greater degree (Anderson et al., 2014). Hence, it would be relevant to take a closer look at how engaging in discourse when preparing for a high-stakes exam can potentially contribute to the students' knowledge formation and eventual exam performance. In this preliminary analysis, Social Network Analysis (SNA) was applied to discussion forum data from students who engaged in an online platform for test preparation and took a standardized exam afterwards. SNA is used in online settings to analyze social relationships and their influence on learning experiences/outcomes. Centrality measures in SNA (degree, closeness, betweenness, eigenvector) were computed for students and analyzed with regard to their subsequent test score (i.e. ACT Composite Score).

2 PRELIMINARY ANALYSIS AND FUTURE WORK

The original dataset included 5,094 students and 19 teachers or administrators (represented as nodes in SNA) who participated in the platform's discussion forums either by posting an original message or replying to an existing one (a total of 11,901 posts). Message posts were found in lessons (52.4%), general questions (24.7%), practice or test questions (21.7%), and live classes (1.2%). Initial inspection showed interactions (i.e., edge/s between students) being heavily skewed towards a single edge (student post had one reply) or no edge at all. Excluding students under this condition and those who did not take a test, the SNA model (directed) in this paper was built from 594 students (with 1,079 unique edges) who either had free access to the platform (fee-waived, e.g. underserved learners) or a purchased access. Among the centrality measures, test performance had a significant correlation with degree centrality (out-degree, $r = 0.24$, $p < 0.05$; in-degree, $r = -0.16$, $p < 0.05$) indicative of better performance being associated with students who reply a lot (out-degree) or students who get replied to a lot (e.g., from an inquiry). In addition, all centrality measures had significant positive correlations (albeit small, ranging from 0.1 to 0.25) with overall duration of online test preparation and practice attempts. Figure 1 shows the visual trends of test performance in relation to their discourse activity from the 594 students in the network. An interesting finding shows fee-waiver students having consistent positive trends in test performance with respect to their centrality measures (except in-degree) suggestive of these students benefitting from increased engagement in discourse (e.g., inquires, responding to others) in their online preparation for a high-stakes exam.

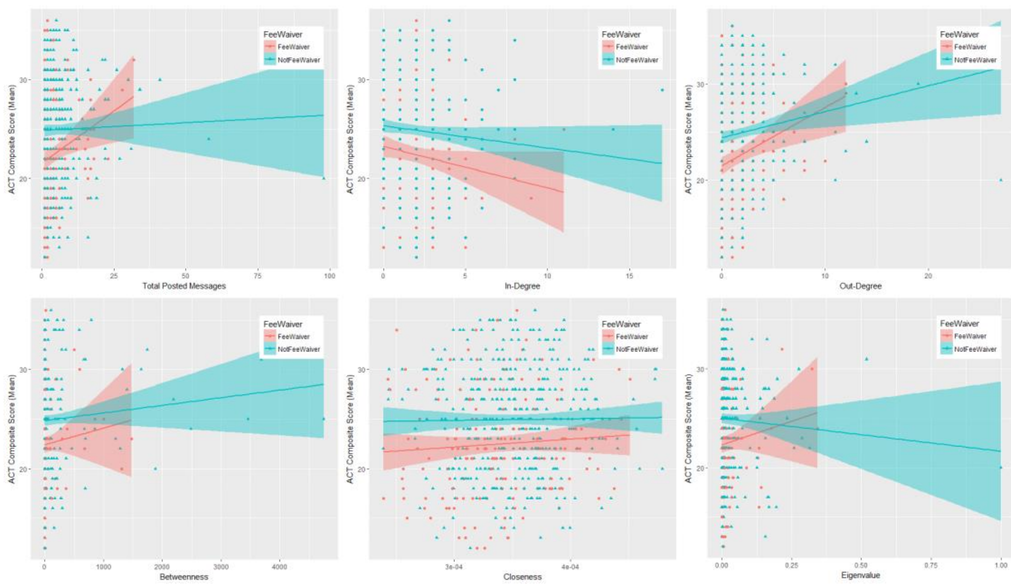


Figure 1: Test Performance and Centrality Measures of Student Discourse during Test Preparation

This SNA-based exploratory analysis has shown how online interaction and discourse between students in the context of test preparation has the potential to aid in test outcomes, in line with research that shows discourse and social relationships during learning being associated with academic performance. Future work in investigating discourse during students' online test preparation include contextualizing these connections through linguistic and semantic analyses, assess together with other preparation activities (i.e., videos or lessons viewed, quizzes and tests, etc.), and analyzing much larger sample of students who participate in discourse for generalizability and impact on practice.

REFERENCES

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses.In *Proceedings of the 23rd Int'l. Conference on World Wide Web* (pp 687- 698).

Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (2015). Investigating How Student's Cognitive Behavior in MOOC Discussion Forums Affect Learning Gains. *International EDM Society*.

Using an Adapted Agile Approach as a Staff and Student Engagement Method for Learning Analytics

Ainsley Hainey
University of Strathclyde
ainsley.hainey@strath.ac.uk

Donna Brawley
University of Strathclyde
donna.brawley@strath.ac.uk

Brian Green
University of Strathclyde
b.green@strath.ac.uk

ABSTRACT: Fostering a culture of openness and transparency is critical to successful implementation of learning analytics. This paper describes the successful use of an adapted Agile project management approach, to build this culture of openness and allow stakeholders to have a voice within the project. This engagement exercise provided an opportunity for stakeholders (students, academic staff, and professional services staff) to participate in the scoping activity of the project and to identify the priorities that would, once developed, have the most impact on their learning, teaching and business needs.

Keywords: Learning analytics, engagement strategy, project management, higher education

1 INTRODUCTION

Agile project management is a flexible and interactive methodology for managing development projects, typically in the areas of engineering and information technology. Agile projects are managed iteratively, with a focus on incremental delivery of product features or functionality and accommodation of changing requirements. The approach requires collaboration between business representatives, customer and supplier.

Studies have shown that when implementing learning analytics, fostering a culture of openness and transparency is critical to its successful implementation (Sclater, Peasgood & Mullen, 2016). This engagement strategy was developed with the aim of addressing learning analytics at a cultural level.

2 METHODOLOGY

The vision for the project was identified through the Learning Analytics Board, the institution's governance of the learning analytics project. This Board consisted of members of Senior Management, Education Enhancement, Information Services, Staff Development Unit, Faculty Representatives, Professional Services representatives, and the Vice President Education from the Students Association. It was agreed that the vision for the project would be implementing learning analytics to address the University's strategic educational priorities. The adapted Agile engagement strategy was developed and this is shown in Figure 1.

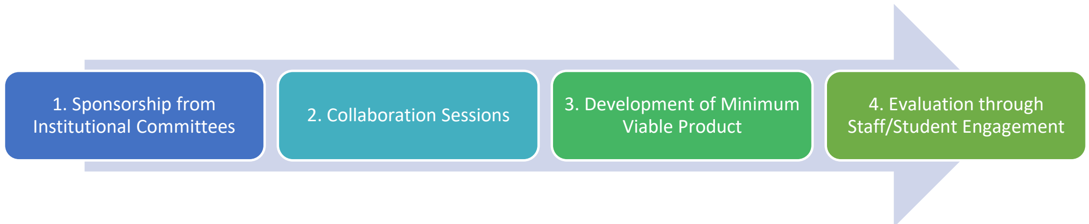


Figure 1: Engagement Strategy

1. Sponsorship from Institutional Committees

Proposals were presented to the Learning Analytics Board, and to Learning Enhancement Committee for approval and endorsement.

2. Collaboration Sessions

Collaboration Sessions were held with staff to scope their requirements for each proposal. This allowed for specifications to be drafted and development work to be undertaken, to implement, in the first instance, a minimum viable product. In addition to the scoping activity, feedback from pilot projects undertaken in phase 1 of the project highlighted potential concerns and issues data and ethics. These concerns were specifically addressed in the collaboration sessions.

3. Development of Minimum Viable Product

A minimum viable product is a version of the new product which allows, with the least effort, the collection of a maximum amount of validated learning. Development work is currently underway in the development of a minimum viable product based on the information gathered at the collaboration sessions.

4. Evaluation through Staff/Student Engagement

Once the minimum viable product has been released, the product will be evaluated (validated learning) by using MoSCoW prioritization by engaging with staff and students. This will allow for Agile development sprints to inform further iterations of the product.

3 CONCLUSIONS

The adapted Agile approach has been invaluable in gathering the views and requirements of staff. It has also had the further impact of identifying elements to help refine the overall institutional strategy for Learning Analytics in enhancing the student experience.

REFERENCES

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education. Retrieved from https://www.jisc.ac.uk/reports/learning-analytics-in-higher-education

The Ties that Support: A mixed methods study of how new direct from high school marginalized and international students create and develop academic and social networks given dispersive institutional landscape

Author(s):
Aurelia Kollasch, Kurt Earnest, K.B. Gwebu
Iowa State University
aureliak@iastate.edu

ABSTRACT: Student engagement outside the classroom often complements student engagement in course-based pursuits. This on-going study explores and the specific structures of academic and social relationships that students, particularly marginalized and international students, create within campus communities. Statistical summaries and social networks analysis provided some insights into student-to-institutional agents and student-to-institutional office networks. Analysis of student-centered data collected showed that student-familial relationships were interdependent and did not impede students' navigation through the transition process. Connections with family and friends outside of the institution as well as with institutional agents can help students envision academic and personal success during transition to college.

Keywords: learning analytics, social network analysis, student-centered engagement

1 INTRODUCTION

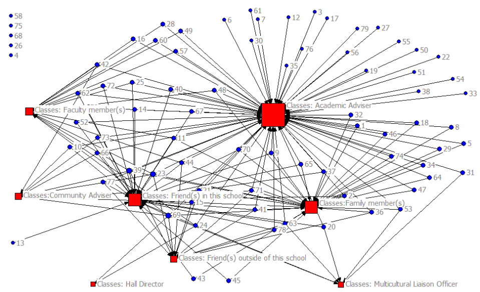
Engagement matters and engagement does not take place in a vacuum rather contexts and environments do matter (Hurtado et al. 1998; Jayakumar, 2008; Locks at al. 2008; Milem and Umbach, 2003). Student engagement outside the classroom often complements student engagement in course-based pursuits. To better understand engagement and integration, educators are asking how to make engagement and integration matter to students. In response, this study focused on the specific character of engagement and the social and academic networks that students develop during their first year of college. Since concepts of integration, involvement, and engagement aim to specify the relationships students have, without actually focusing on the structure of students' campus connections, our research study used a mixed-methods design of social network analysis (SNA) and interviews to provide the conceptual and analytical tools to illuminate or highlight the patterns of ties among students. Residence halls served as living spaces conducive to social and academic interaction among students, and therefore, offered an essential research setting in the dispersive institutional landscape.

2 METHODS AND INITIAL FINDINGS

A sociometric questionnaire that included five major questions asking about students' individual and institutional academic, professional, and personal networks was administered to over 200 multicultural and international students who resided in designated residence halls at the Midwestern University in October 2017. The questionnaire yielded 79 valid responses from 39 (49.4%) females and 40 (50.6%) males. 88.6% of respondents were multicultural students when only 11.4% comprised the international sample. In addition, open-ended interviews were conducted in

December 2017 to seek further articulation of students' individual and institutional networks during their transition to college life.

The resulting social networks varied greatly in size, diversity, and the extent to which an individual's ties were connected to each other. Both international and multicultural students found areas on campus where they developed academic and professional relationships and fostered positive connections with institutional agents (Figure 1 and 2).



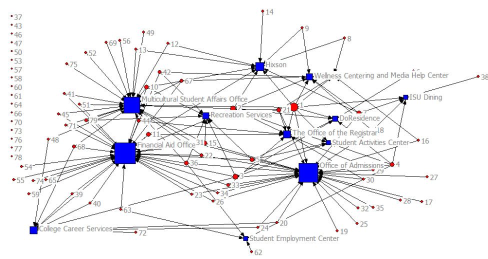


Figure 1: Students' academic engagement networks with institutional agents.

Figure 2: Students' academic engagement networks with offices on campus.

Student's family and friends outside of the institution were strong ties while institutional agents represented weak ties. The strength of ties was defined by the frequency of contact students reported having with various members of their academic, personal, professional networks. Because students were asked how often they seek academic, personal, and professional advice from individuals and institutional resources, strong ties were those relationships that occurred with more frequency, while weak ties were those relations that respondents relied on with less frequency. The graphs portrayed a static description of students' networks. However, individual interviews revealed that students relied on different networks when experiencing different challenges in their transition to college life. Particularly, institutional agents such as academic advisors played a more prominent role in these networks when students needed help in the academic area and interactions occurred at least once a week. However, academic networks seemed to be small and less diverse during the transition period where many respondents relied on their family and friends outside of the current institution to manage academic transition. Women of color were least connected to faculty and staff around campus, but their familial ties were the strongest. On the other hand, the institutional networks were more diverse showing that students tried to navigate the new community by visiting certain offices on campus for a particular type of assistance, with the Office of Financial Aid and the Office of Admissions being visited at least once a week during the transition time.

Even though ties with families and friends outside the institution may be strong in nature , they may not expose the student to diverse sources of information. The familial and peer networks were crucial sources of support; therefore, intervention planning should focus on outreaching to family members with information that may assist both students and families in navigating college life. Informational capital was a necessary form of social capital during the transition process. This also lends support to findings that many students did not need to separate from their external communities to feel a sense of belonging on campus. Rather, these student-familial relationships can be considered interdependent and not impeding on their navigation during the transition process. While connections with family and friends as well as with institutional agents can help students envision success in higher education, the informational capital may be interconnected and an interrelated component of educational experience along with integration to further capture sense of belonging and community on campus.

REFERENCES

Hurtado, S., Milem, J., Clayton-Pederson, A. R., & Allen, W. R. (1998). Enhancing campus climates for racial/ethnic diversity: Educational policy and practice. *The Review of Higher Education*, *21*(3), 279–302.

Jayakumar, U. M. (2008). Can higher education meet the needs of an increasingly diverse and global society? *Harvard Education Review*, *78*(4), 615–651.

Locks, A. M., Hurtado, S., Bowman, N. A., & Oseguera, L. (2008). Extending notions of campus climate and diversity to students' transitions to college. *Review of Higher Education, 31*(3), 257–285.

Milem, J. F., & Umbach, P. T. (2003). The influence of pre-college factors on predispositions regarding diversity activities in college. *Journal of College Student Development, 44*(5), 611–624.

Doctoral Consortium

"How does it work for you?": Exploring the differential effects of student-facing learning analytics applications

Author: Lisa-Angelique Lim
University of South Australia
lisa.lim@unisa.edu.au

ABSTRACT: Recent developments in educational technologies have provided a viable solution to the challenges associated with scaling feedback to all students in large courses. However, there is currently little empirical evidence about the impact such scaled feedback has on student learning. This doctoral research will evaluate the impact of novel learning analytics reporting processes to identify how feedback can support and develop students' self-regulated learning proficiency. The impact of this will be evaluated with respect to different groups of learners, especially in terms of learners' emotions and self-efficacy, which, though recognised as important factors in self-regulation, is a significant research gap in the field of LA. This research will build on and contribute to the theory of motivation, metacognition, and self-regulated learning, as well as contribute to the area of pedagogical analytics by documenting how different learners respond to personalised, LA-based feedback.

Keywords: Personalised feedback, large enrolment courses, self-regulated learning, differential impact

1 BACKGROUND

A challenge for contemporary educators is how to provide feedback to all students in large courses in a personalised, timely and instructive manner. The development of technology-based, personalised feedback to students is a significant innovation for education: this innovative approach positions "one of the most influential aspects in the quality of the student learning experience, feedback, within the current research space of the EDM [educational data mining] and LA communities" (Pardo, Poquet, Martinez-Maldonado, & Dawson, 2017, p.168). Learning analytics provides the opportunity to deal with this challenge by automating the collection of data related to students' learning and performance and transforming these into useful metrics that can be fed back to every student in a course, either through visual dashboards, recommender systems, personal emails, or phone calls from the instructor. The automated personalised feedback system aims to help students make evidence-based decisions about their learning, and to foster self-regulated learning, a critical 21st century skill.

Although the number of student-facing LA applications is on the rise, there is no conclusive research to inform whether existing student facing LA applications are effective in improving learning outcomes or motivation (Bodily & Verbert, 2017; Ferguson et al., 2016). The research is not conclusive as to whether closing the loop for students in this way actually improves their learning, or whether such feedback provisions work equally well for all students. The latter is an important issue, since in today's higher education environment, there is a growing and increasingly diverse, student population.

This doctoral research aims to explore the differential effects of personalised feedback based on LA. In particular, it aims to document how individual differences and disciplinary backgrounds affect how students interpret and respond to LA feedback.

1.1 Supporting learners with data-driven personalized feedback and factors affecting feedback efficacy

The literature abounds with examples of institutionally-adopted data-driven feedback systems: E2 Coach at the University of Michigan (Wright, McKay, Hershock, Miller, & Tritz, 2014), the Competency Map at Capella University (Grann & Bushway, 2014), the Student Relationship Engagement System, SRES at the University of Sydney (Liu, Bartimote-Aufflick, Pardo, & Bridgeman, 2017), or the OnTask Project (Pardo, Martínez-Maldonado, et al., 2017). At the same time, there has been criticism that LA approaches have been too reliant on generic learner log data in LMSs and student demographic and educational background information from institutional student information systems to generate early-warning systems (Gašević, Dawson, Rogers, & Gasevic, 2016). Although these may be considered data-driven and therefore empirically based, employing such a generic approach across all courses may not be sufficient to help students in terms of providing actionable intelligence (Wise, 2014) to improve their learning. Hence, more contextualised LA-based approaches to feedback and support are recommended by practitioners (Dawson, Jovanovic, Gašević, & Pardo, 2017; Liu et al., 2017). As shown by Gašević et al (Gašević et al., 2016), the instructional conditions in a course cannot be neglected in LA-based interventions.

That feedback is central to learning has been widely established (Hattie, 2014; Hattie & Timperley, 2007; Kluger, Denisi, & Steinberg, 1996). In terms of Winne and Hadwin's (Winne & Hadwin, 1998) COPES model of self-regulated learning, students set goals, and monitor how their learning strategies are progressing toward those goals. This monitoring provides an internal feedback loop that relies on both internal and external feedback to help students regulate their learning. Feedback affects learners' evaluation of the products of their learning and effectiveness of study tactics and strategies used. Such evaluation raises students' awareness of how they are learning (i.e., monitoring) and whether they are on the right track. As well, evaluation helps them to know how to adjust their learning strategies to reach learning goals, thereby leading to enhanced achievement.

The effects of feedback have been found to be moderated by individual differences (Winstone, Nash, Parker, & Rowntree, 2017). In particular, prior academic achievement may influence the effectiveness of the feedback. In an experimental study (Teasley, 2017) where students were presented with hypothetical Blackboard dashboards as feedback, students with higher GPA reported that they would engage less with the LA-based feedback, as compared to their lower-performing peers. On the other hand, students with lower GPA reported that they would find the recommended follow-up actions more helpful. While the Blackboard study was presented in a hypothetical setting, few studies have been carried out in live courses, investigating how prior academic achievement might affect the way students engage with LA-based feedback systems.

Furthermore, feedback efficacy may be moderated by students' motivation for their learning the subject: according to Biggs (Biggs, 1999) students who harbour a more surface or extrinsic motivation for learning in a course will adopt surface strategies such as rote learning and memorisation, while students with deep or intrinsic motives for learning the course will adopt deep strategies which deepen understanding of the subject. When it comes to feedback then, students with more surface approaches to learning may choose not to respond to the feedback as their main goal is to simply pass the course; in contrast students with deeper approaches to learning may take actionable intelligence from the feedback in order to enhance their learning. Thus far, this hypothesis has not been tested with respect to technologically-delivered feedback based on LA.

1.2 Self-regulated learning: the role of self-efficacy and emotions

The main goal of providing personalised feedback through LA is to foster students' self-regulated learning (SRL). SRL has been studied extensively and found to be positively associated with academic success (Dent & Koenka, 2016; Pintrich & De Groot, 1990). Students proficient in time management, metacognitive awareness, critical thinking and perseverance have been found to be more successful academically (Broadbent & Poon, 2015). As noted by Roll and Winne (2015), "Learning analytics are reports of analyses of data that describe features of, and factors that influence, SRL" (p.8). A clear understanding of SRL is vital for evaluating LA-based feedback interventions in order to evaluate impact on this outcome.

Theories of SRL (Mega, Ronconi, & De Beni, 2014; Pintrich & De Groot, 1990; Winne & Hadwin, 1998; Zimmerman, 2000) have featured self-efficacy as an important motivational influence in self-regulated learning behaviour. Self-efficacy, in particular academic self-efficacy, refers to one's personal belief about being able to bring about desired learning goals, and is therefore a component of motivation for learning. It is argued that for effective self-regulation, learners need to have a strong belief in their ability to regulate their learning and to perform well (Pintrich & De Groot, 1990). Self-efficacy has also been shown in at least three meta-reviews (reported in Broadbent, 2016) to have a motivating influence on academic performance. Although the role of self-efficacy has been hypothesised as a variable in LA research (Gašević, Dawson, & Siemens, 2015), this concept has so far received little empirical investigation.

The last decade has seen an interest in the role of affect (emotions) in learning. In any learning situation, students experience a variety of affective states which can either positively or negatively impact self-regulated learning and ultimately, learning outcomes (D'Mello, 2013). There is particular interest in studying how affect shapes student engagement and learning (Linnenbrink-Garcia & Pekrun, 2011). Affect has been found to be an important motivational factor in learning (Pekrun, 2006; Webster & Hadwin, 2015), and as one indicator of student engagement (Fredricks, Blumenfeld, & Paris, 2004). There is also a body of research in affective computing, exploring the detection of learners' affective states during technologically-mediated learning (D'Mello & Graesser, 2014). Although the availability of data and advances in technology have provided more opportunity to understand affect in learning, little research thus far has examined how LA interventions impact learner affect and subsequently their motivation—particularly self-efficacy—and self-regulated learning.

2 AIM OF THE RESEARCH

This doctoral research builds on existing LA research by examining differential impact of personalised feedback based on LA. This research focuses on the students' perspective of LA, a much-needed perspective for personalised learner support (Roberts, Howell, Seaman, & Gibson, 2016). While other research has drawn on hypothetical data to evaluate impact of student-facing LA interventions, this research will be carried out 'in the field', with data collected from authentic settings where the intervention is being deployed. Moreover, this research will employ a rigorous evaluation framework based on SRL, relying not only on student perceptions of their experience with the LA intervention, but backing these with analysis of impact on learning behaviours and academic performance. The following research questions will be addressed: 1) How do students interpret their LA-based feedback? 2) What is the impact of LA feedback on (a) students' affect, (b) students' self-efficacy, and (c) students' online learning behaviours? 3) Does the effect of LA-based feedback on a) self-efficacy and b) academic performance in the course differ between students with different approaches to learning and academic disciplinary background?

3 METHODOLOGY

This research will comprise sub-studies of the impact of personalized feedback based on LA in different disciplines and institutions. Each sub-study will follow a similar protocol. To address RQ1, qualitative data will be collected through interviews. These will be designed to draw out what students understood from their feedback, how the feedback affected their motivation for learning, and what actions they planned to take in response to the feedback. To address RQ2, students' affective response to the feedback will be measured using Russell and colleagues' (Russell, Weiss, Mendelsohn, & Sarason, 1989) affect grid. Baseline and end-of-course quantitative data on students' approach to learning using the R-SPQ-2F (Biggs, Kember, & Leung, 2001) and self-efficacy subscale of the MSLQ (Duncan & Mckeachie, 2005) will be collected for this question, and learner trace data from course LMS will be examined. To answer RQ3, approach to learning data will be analysed using cluster analysis to identify distinct groups, and analyses of variance will be conducted on quantitative data to explore differential impact.

4 CURRENT STATUS & RESULTS ACHIEVED

To date, the first sub-study has been conducted, and a report submitted as a full paper to LAK18. This first investigation was carried out in a large, first-year undergraduate course in the health sciences. Log data from the LMS and e-book, as well as performance on course assessments, from three years of course offerings (n_{2015}=265, n_{2016}=277, n_{2017}=242) were analysed. The latest course offering (in 2017) involved the personalised, LA-based feedback support intervention. The results indicated that the intervention cohort showed significantly different patterns of online engagement in line with the recommendations of the feedback and performed better in terms of final grades. In particular, the feedback emphasised the importance of regular engagement with the e-book[1] to optimise course performance; as compared to previous cohorts which showed 'cramming' behaviour, engagement with the e-book was sustained until

[1] http://www.mheducation.com/highered/platforms/connect.html

the end of the semester. In addition, differential impact on quiz and course marks was found, with the third highest quartile group of prior achievement deriving the most benefit from the feedback. Impact on affective response and academic performance did not differ by approach to learning.

For the next steps, interviews with students will be conducted to further understand students' experience of the feedback support based on LA, how students made sense of it, and how it affected their motivation and learning in the course. More in-depth analysis of the online engagement data from the intervention cohort will also be conducted using EDM, to investigate specifically how students' behavior changed in response to the feedback. Finally, further sub-studies are being planned to find answers to the research questions.

REFERENCES

Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher education research & development, 18*(1), 57-75.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The Revised Two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal Of Educational Psychology, 71*, 133-149.

Bodily, R., & Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. *Proceedings of the 7th International Conference in Learning Analytics and Knowledge*, ACM. doi:10.1145/3027385.3027403

Broadbent, J. (2016). Academic success is about self-efficacy rather than frequency of use of the learning management system. *Australasian Journal of Educational Technology, 32*(4).

Broadbent, J., & Poon, W. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1-13. doi:10.1016/j.heduc.2015.04.007

D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology, 105*(4), 1082-1099. doi:10.1037/a0032674

D'Mello, S. K., & Graesser, A. C. (2014). Feeling, thinking, and computing with affect-aware learning. *The Oxford Handbook of Affective Computing*, 419-434.

Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017). From prediction to impact: Evaluation of a learning analytics retention program *Proceedings of the 7th International Conference on Learning Analytics and Knowledge*. ACM.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*(3), 425-474. doi:10.1007/s10648-015-9320-8

Duncan, T. G., & Mckeachie, W. J. (2005). The making of the Motivated Strategies for Learning Questionnaire. *Educational Psychologist, 40*(2), 117-128. doi:10.1207/s15326985ep4002_6

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., . . . Vuorikari, R. (2016). *Research evidence on the use of learning analytics: Implications for education policy*. Joint Research Centre, Seville, Spain.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*(Supplement C), 68-84. doi:https://doi.org/10.1016/j.iheduc.2015.10.002

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71.

Grann, J., & Bushway, D. (2014). Competency map: Visualizing student learning to promote student success. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, 168-172. ACM. doi:10.1145/2567574.2567622

Hattie, J. (2014). *Visible learning for teachers: Maximizing impact on learning*. Florence, GB: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi:10.3102/003465430298487

Kluger, A. N., Denisi, A., & Steinberg, R. J. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284. doi:10.1037/0033-2909.119.2.254

Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' emotions and academic engagement: Introduction to the special issue. *Contemporary Educational Psychology, 36*(1), 1-3. doi:10.1016/j.cedpsych.2010.11.004

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-driven personalization of student learning support in higher education. In A. Peña-Ayala (Ed). *Learning Analytics: Fundaments, Applications, and Trends*, 143-169.

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-refulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*(1), 121-131. doi:10.1037/a0033546

Pardo, A., Martínez-Maldonado, R., Buckingham Shum, S., Schulte, J., McIntyre, S., Gašević, D., . . . Siemens, G. (2017). Connecting data with student support actions in a course: A hands-on tutorial. In. *Proceedings of the Seventh International Conference on Learning Analytics & Knowledge*, 522-523. ACM. doi:10.1145/3027385.3029441

Pardo, A., Poquet, O., Martinez-Maldonado, R., & Dawson, S. (2017). *Provision of data-driven student feedback in LA and EDM*. Paper presented at the Handbook of Learning Analytics.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student Attitudes toward Learning Analytics in Higher Education: "The Fitbit Version of the Learning World". *Frontiers in Psychology, 7*(1959). doi:10.3389/fpsyg.2016.01959

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics, 2*(1), 7-12.

Russell, J. A., Weiss, A., Mendelsohn, G. A., & Sarason, I. G. (1989). Affect Grid: A Single-Item Scale of Pleasure and Arousal. *Journal of Personality and Social Psychology, 57*(3), 493-502. doi:10.1037/0022-3514.57.3.493

Teasley, S. D. (2017). Student Facing Dashboards: One Size Fits All? *Technology, Knowledge and Learning*, 1-8. doi:10.1007/s10758-017-9314-3

Webster, E. A., & Hadwin, A. F. (2015). Emotions and emotion regulation in undergraduate studying: Examining students' reports from a self-regulated learning perspective. *Educational Psychology, 35*(7), 94-818. doi:10.1080/01443410.2014.895292

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, A. C. Graesser (Eds). *Metacognition in educational theory and practice*, 27-30. Mahwah, NJ : Lawrence Erlbaum Associates.

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17-37. doi:10.1080/00461520.2016.1207538

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge*. doi:10.1145/2567574.2567588

Wright, M. C., McKay, T., Hershock, C., Miller, K., & Tritz, J. (2014). Better than expected: Using learning analytics to promote student success in gateway science. *Change: The Magazine of Higher Learning, 46*(1), 28-34. doi:10.1080/00091383.2014.867209

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82-91. doi:10.1006/ceps.1999.1016

Developing a Learning Analytics Intervention Design and tool for Writing Instruction

Antonette Shibani
University of Technology Sydney
antonette.aileenshibani@student.uts.edu.au

ABSTRACT: Academic writing can be supported by the provision of formative feedback in many forms including instructor feedback, peer feedback and automated feedback. However, for these feedback types to be effective, they should be applied in well-designed pedagogic contexts. In my pilot study, automated feedback from a writing analytics tool has been implemented in pedagogic interventions, which integrate learning analytics technologies to existing practice in an educational context. To improve the learning design and to study the use of human insights in this context, a peer discussion module is planned to be added. This kind of peer discussion can augment automated feedback applications by making students aware of the limitations of such artificial intelligence powered feedback, and develop writing literacy by providing additional contextual feedback for their peers. The learning analytics intervention design when tested across different disciplines can validate the usefulness of this approach to improve students' academic writing in authentic pedagogic contexts. The design can be implemented using a learning analytics tool which is developed to facilitate the intervention and provide analytic capabilities by collecting learner data.

Keywords: academic writing, automated feedback, peer feedback, learning design, learning analytics

1. INTRODUCTION

Academic writing is challenging to learn for many students, who could be supported by the provision of formative feedback on the writing. Formative feedback aids students to gain awareness on where they stand in terms of goals in their current work and how to improve their progress by determining the way forward (Sadler, 1989). Formative feedback on students' drafts can help them to improve their writing through the application of the feedback in revisions of their text. Due to the time-consuming nature of instructor provided feedback for student drafts, especially in large cohorts, alternative forms of feedback like peer feedback and automated feedback have been studied to help improve students' writing skills.

Automated tools use computational techniques to provide immediate feedback on students' drafts. They reduce waiting time and human effort, ensure consistency and encourage students to practise writing and revision by providing feedback multiple times while writing drafts (Shermis, Raymat, & Barrera, 2003). Such writing analytics tools which offer the potential to provide timely automated feedback on students' writing are good examples of learning analytics applications in pedagogy. However, these technologies have to be embedded in the curriculum by implementing them in well-designed contexts for better uptake by students, and for solving existing pedagogical issues using learning analytics. This move from developing learning analytics technologies to integrating them as part of a larger educational context can be done using 'pedagogical learning analytics intervention design' which refers to the "systematic efforts to incorporate the use of analytics as a productive part of teaching and learning practices in a given educational context" (Wise, 2014). This adds value to learning by closing the gap between potential and actual use of technologies. The alignment of learning analytics to learning design also provides a contextual framework to document the pedagogic intent of analytics applications and to collect data for its evidence (Lockyer, Heathcote, & Dawson, 2013).

While working with learning analytics tools and dashboards, the human context is often emphasized as central in interpreting and making sense of the analytics (Siemens, 2012). This is because learning is a complex activity involving social processes. Sense-making and interpretations are hence important in writing analytics tools, in order for students to understand and implement the automated feedback that is provided on their writing. One way of providing sense-making support is through peer feedback and discussion, where students can interpret, discuss and critique automated feedback on writing with their peers. This approach of combining peer discussion and automated feedback has two core benefits. First, peer discussion overcomes limitations in automated feedback by complementing it with contextual feedback by peers to capture features missed by the tool. This brings in a human context which is lacking in automated feedback and enhances the social and cognitive processes involved in writing. Students also learn from each other while providing feedback on each other's writing by making judgements about their performance (Allal, Lopez, Lehraus, & Forget, 2005). Second, automated feedback may address a concern in peer feedback regarding student's abilities to provide meaningful feedback, by scaffolding this feedback and provoking discussion around the identified features.

I propose a design that combines known effective practices like peer feedback and discussion with automated feedback, to complement each other. The design will be applicable in pedagogic contexts where learning analytics can augment existing learning designs to improve students' writing skills. To represent the design in a theoretical and practical way for implantation in practice, abstractions like conjecture mapping (Sandoval, 2014) and design patterns (Goodyear, 2005) will be developed. Thus, the main aim of my research is to develop a pedagogic learning analytics intervention design that combines automated and peer feedback to improve students' academic writing, and design patterns to help implement it in classroom settings across different pedagogic contexts. In doing that, I will study the following as part of my research:

- The impact of automated feedback in student writing
- The impact of the inclusion of a peer discussion component with automated feedback
- The implementation of learning design across different contexts

2. CURRENT STATUS OF WORK

In the first part of my doctoral study, a learning design aligned with learning analytics was developed by embedding the automated feedback tool Academic Writing Analytics (AWA) in a law subject (Shibani, Knight, Buckingham Shum, & Ryan, 2017). This was to help students understand the role of rhetorical structures and the usage of automated feedback on them in a way that it can be applied for their subject essay writing. These rhetorical structures guide the reader through the argument structure of a text, and are a key part of academic writing (Hyland, 2005), with their presence having some (small) relationship to essay quality (Simsek et al., 2015). An intervention grounded by pedagogy was designed for students to learn essay writing and revision skills based on rhetorical moves in the context of their subject curriculum by augmenting existing practice with learning analytics. The intervention consisted of a set of tasks which were completed by the students individually in a classroom during a tutorial session. To carry out the pedagogical intervention, I developed a tool called AWA-Tutor as an extension of AWA which guides students through different activities to learn rhetorical writing (Shibani, 2018). The tool also collects data when guiding the students through the tasks to enable learning analytics to collect evidence for the learning design. The design consisting of the tasks and the data collected are listed in Table 1 in the Appendix.

From this intervention, the impact of the types of feedback provided was studied using the essays revised by students of different assigned groups, and by automated feedback group in particular (Shibani et al., 2017). It was because this area studying the impact on students' writing has not been researched extensively in many tools that provide automated feedback, since the focus has mainly been on the accuracy of such tools. The analytic data collected from the tool also enabled the analysis of students' revision process which was previously hard to track. Studying this revision data and processes in the context of student essays aids to gain insights on processes involved in student writing (Shibani, Knight, & Shum, 2018, Submitted for Review).

Thus, the writing analytics tool was embedded within a curriculum as a pedagogic intervention for improving students' writing ability by making use of several tasks. This design provided an intervention design for students to learn writing skills, and a platform to deliver the activity, with integrated learning analytics for collecting data for instructors and researchers, and delivering instant feedback to students. All the tasks in this intervention were completed by students individually. The next design will introduce a peer discussion component along with automated feedback to study its effectiveness (Shibani, 2017). This design has been piloted, with some improvements made in the design and in the tool. The design included a new component at the beginning of the activity which explains the use of rhetorical moves and discourse markers to students. This was done by providing a supporting material which was read aloud during the tutorial. The upgraded tool included descriptions of rhetorical moves identified by the automated feedback component with example sentences. Preliminary analysis from this data has shown increased acceptance and understanding of the activity among students, although there is no significant different noted across the conditions. The next study will study the effects in detail based on a design that encompasses automated feedback and peer discussion at student level to study their differences, with some improvements made to automated feedback as well. The learning design is improved across iterations, based on observations from the previous implementation and feedback from the instructor and students. When it is stabilized for wide usage, a design package consisting of the design patterns and a tool to help implement the design in pedagogical contexts will be developed for practitioners adopting this writing instruction approach.

3. GOALS

The main goal of my research is the development of a learning design, abstractions and a tool that can help implement the design to be used by researchers and practitioners for writing instruction. The design will also be tested in a discipline other than the one it was originally developed for to validate its transferability. The abstractions and learning patterns that evolve can be used by practitioners for implementing the design in their classroom. Using this design, I also aim to study some research questions which are not extensively studied in past literature.

3.1 Research Questions

The overall aim of my research is to develop an effective learning analytics intervention design that can be used across disciplines to study and impact students' writing in authentic contexts by making use of both automated and peer feedback. The specific research questions are:

1 What is the impact of automated feedback on student writing?

- What are students' perceptions of automated feedback?
- What is the impact of automated feedback on student revisions?

2 What is the impact of automated feedback when combined with peer discussion on student writing and revisions?

- Do students produce higher quality texts when a peer feedback component is added to automated feedback?
- How do peer discussion dynamics impact the outcome?
- What kinds of automated and peer feedback did students act on?
- What is the student self-reported value of peer feedback in combination with automated writing feedback?

3 How does the transfer of intervention design work across disciplines?

- What are the abstractions to be developed to help a practitioner implement the learning analytics intervention design in their discipline?
- What is the student self-reported value of the intervention in the other discipline?

3.2 Methodology

The overarching methodology used in this research is design-based research (DBR), which is "a systematic, but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in real world settings, and leading to contextually-sensitive design principles and theories'' (pp. 6–7) (Wang & Hannafin, 2005). Changes are made to the design of the intervention and the tool based on the feedback from the earlier implementation and discussion with the instructor until the design is stabilized.

From the pilot studies which were carried out earlier, further improvements that can be made to automated feedback format were identified for better student uptake. New ways of helping students by providing contextual feedback and improving social sense making in writing were also trialed by including peer discussion in the design. The proposed design will hence be an extension of the previous studies, by making changes to the tool and design in addition to the inclusion of a peer discussion component, as shown in Figure 1.

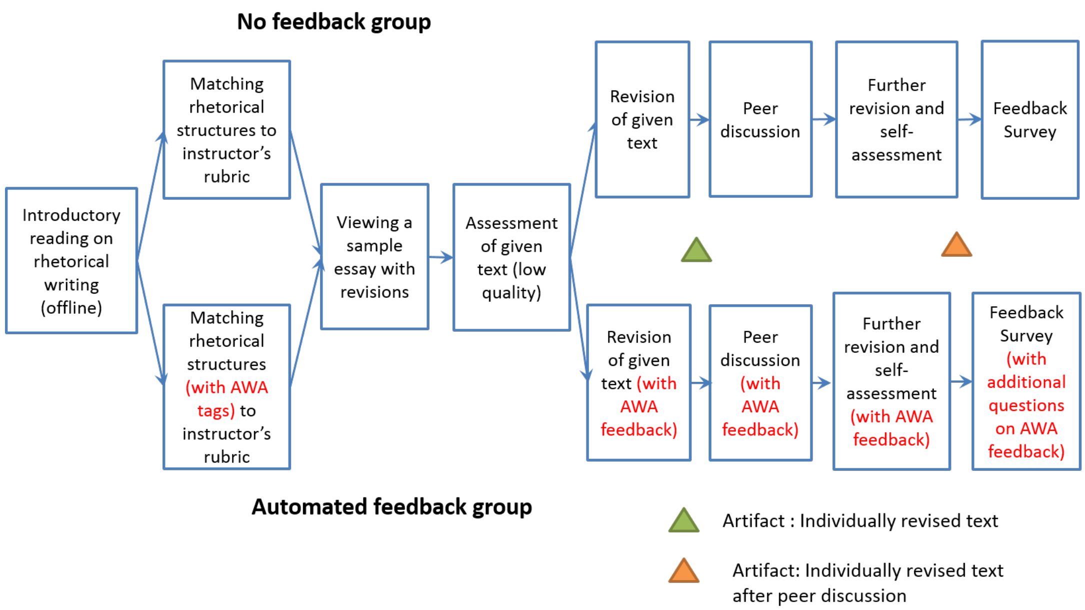


Figure 1: Proposed learning analytics intervention design for rhetorical writing instruction

This design will be used to study the impact of automated feedback and dyadic peer discussion on students' writing by comparing individually revised texts and revised texts after peer discussion in the context of with/without automated feedback. The peer discussion dynamics which is likely to have an effect on the outcomes will be studied using qualitative analysis building on the work of discourse-centric learning analytics (Knight & Littleton, 2015). Based on the results of the effect of peer discussion component, the design may be updated in the future to run with/ without peer discussion.

The current design is for improving rhetorical writing of students by teaching the structure of text in terms of rhetorical structures, discourse markers and automated feedback on them featuring the use of AWA tool. However, it can also be potentially applied to writing instruction by making use of tools that provide other types of feedback E.g. cohesion. The design is also pedagogically sound to implement even without the use of automated feedback and analytics capabilities.

4. CONCLUSION

Academic writing in students can be supported by the use of resources and tools in most pedagogic contexts. My doctoral research hence focusses on developing an efficient learning design by combining automated and peer feedback for writing instruction. A learning analytics pedagogic intervention design and automated tools to help carry out the intervention should enable new ways to embed learning analytics applications in authentic contexts, and ultimately, improve writing. The validated design could be potentially transferrable across contexts with the development of standardized abstractions.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Prof. Simon Buckingham Shum and Dr. Simon Knight for their constant support and guidance in my research.

REFERENCES

Allal, L., Lopez, L. M., Lehraus, K., & Forget, A. (2005). Whole-class and peer interaction in an activity of writing and revision *Writing in context (s)* (pp. 69-91): Springer.

Goodyear, P. (2005). Educational design and networked learning: Patterns, pattern languages and design practice. *Australasian Journal of Educational Technology, 21*(1).

Hyland, K. (2005). *Metadiscourse*: Wiley Online Library.

Knight, S., & Littleton, K. (2015). Discourse-centric learning analytics: mapping the terrain. *Journal of Learning Analytics, 2*(1), 185-209.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439-1459.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science, 18*(2), 119-144.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the learning sciences, 23*(1), 18-36.

Shermis, M. D., Raymat, M. V., & Barrera, F. (2003). Assessing Writing through the Curriculum with Automated Essay Scoring. *ERIC document reproduction service no ED 477 929*.

Shibani, A. (2017). Combining automated and peer feedback for effective learning design in writing practices. In W. Chen et al. (Ed.). *Proceedings of the 25th International Conference on Computers in Education*, New Zealand.

Shibani, A. (2018). AWA-Tutor: A Platform to Ground Automated Writing Feedback in Robust Learning Design. *Proceedings of the 8th International Conference on Learning Analytics & Knowledge (LAK18)*, Sydney.

Shibani, A., Knight, S., Buckingham Shum, S., & Ryan, P. (2017). Design and Implementation of a Pedagogic Intervention Using Writing Analytics. In W. Chen et al. (Ed.). *Proceedings of the 25th International Conference on Computers in Education*, New Zealand.

Shibani, A., Knight, S., & Shum, S. B. (2018, Submitted for Review). Understanding Students’ Revisions in Writing: From Word Counts to the Revision Graph.

Siemens, G. (2012). Sensemaking: Beyond analytics as a technical activity. *Presentation at the EDUCAUSE ELI*.

Simsek, D., Sándor, Á., Buckingham Shum, S., Ferguson, R., De Liddo, A., & Whitelock, D. (2015). Correlations between automated rhetorical analysis and tutors' grades on student essays. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 355-359), NewYork.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational technology research and development, 53*(4), 5-23.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. *Proceedings of the Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* (pp. 203-211).

Appendix

Table 1: Design of tasks in iteration 1

Task	Purpose	Tool Design	Data collected
Matching exercise	Enabling students' understanding of the instructor's rubric elements (Assessment criteria) for assessing rhetorical essays by matching sample rhetorical moves that correspond to the rubric elements.	Interactive drag and drop interface (augmented by analytics for immediate feedback)	Time taken to complete the task
Viewing exemplar revisions	Enabling learning to revise using exemplars and provide understanding of the current activity requirements by giving students a sample revised essay with changes made by the instructor to improve it highlighted.	Displaying a tracked version of a sample improved essay in pdf format that shows revisions made using rhetorical moves and their rationale.	Nil
Essay Assessment	Assessing the given low quality essay by understanding the assessment criteria, and identifying possible revisions that can be made to improve its quality.	Guiding questions to enable assessment and reflection.	Student responses
Revision & Self-assessment (Main task)	Practising revision by the assessed essay to improve its quality using feedback (if provided, based on the group assignment). In this first iteration, students were assigned to one of the following feedback groups to study the feedback effects: AWA Feedback Group, Instructor Feedback Group and No Feedback Group.	A revision editor for all groups and an interface to receive automated feedback for AWA feedback group when requested (augmented by analytics for immediate feedback)	Drafts during the revision process, Revised essays, Student responses for assessment questions, Requests for automated feedback
Task evaluation	Receiving feedback from the students on the whole activity for evaluation of the task design. Students also get to download an instructor's sample revised essay and their own revised essay for reflection.	Evaluation questions to provide feedback	Student responses, Click history tracked to record the download of revised essays.

Adaptive learning analytics dashboard for self-regulated learning

Ioana Jivet
Open University of the Netherlands
ioana.jivet@ou.nl

ABSTRACT: Learners need to self-regulate their learning in order to be successful in an online learning environment. However, existing platforms offer little support for the development of such skills. Learning dashboards are learning analytics interfaces built with the purpose of making learners aware of their learning performance and behaviour and supporting self-reflection. However, most of the existing dashboards follow a "one size fits all" philosophy disregarding individual differences between learners. Moreover, there is a strong emphasis on comparison and competition with peers when reporting data back to the learners through dashboards, posing a long-term threat that "being better than your peers" becomes the definition of a successful learner. Throughout this PhD project, we aim to research, develop, implement and evaluate a learning dashboard that caters to the individual needs of learners throughout the self-regulated learning process.

Keywords: Learning analytics; personalization; dashboards; online learning; self-regulated learning, achievement goal orientation.

1 INTRODUCTION

Learners with self-regulated learning (SRL) skills are metacognitively, motivationally and behaviourally active participants in their own learning (Zimmerman, 1990). Such learners plan, set goals, self-monitor and self-evaluate their knowledge acquisition process. Additionally, they create environments that optimise learning by seeking out advice, information and places where they are most likely to learn. Studies of SRL in the traditional classroom have shown that students who have highly developed SRL skills also have higher academic achievements (Pintrich & De Groot, 1990), yet students do not receive much explicit instruction in schools on how to effectively develop their SRL skills (Winne, 1996). Moreover, as online learning environments become increasingly common, the distance between learners and teachers grows and thus, the importance of SRL becomes crucial (Jansen et al., 2017). Online learning environments also provide opportunities for supporting learners in acquiring SRL skills due to the availability of data generated by these environments that can be presented back to the learners as feedback.

One such opportunity presents itself through *learning analytics dashboards*. Learning Analytics (LA) aims to exploit the potential of the increasingly large amounts of data describing interactions, personal data and academic information generated by the widespread use of online learning environments (Ferguson, 2012). As defined by Schwendimann et al. (2016), learning analytics dashboards are "single displays that aggregate different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualisations". Although the research in the field of learning analytics is relatively new and promising, a series of shortcomings overshadow the potential of using LA dashboards to support students in their self-regulated learning.

Firstly, although learning dashboards could be used as powerful metacognitive tools for learners (Charleer, Klerkx, Duval, De Laet, & Verbert, 2016), very few dashboards are targeted at learners (Schwendimann et al., 2016). Secondly, although a number of scholars argue that learning analytics research should be deeply grounded in learning sciences (Lonn, Aguilar & Teasley, 2015), there is a much stronger focus on the ”analytics” aspect and less emphasis on the “learning” (Gašević, Dawson & Siemens, 2015). Moreover, LA should be seen as an educational approach guided by pedagogy and not the other way around (Greller & Drachsler, 2012). Thirdly, most dashboards are designed following the presumption that “one size fits all” without considering the individual differences between learners. Research in the field of learning science indicates that both external factors, i.e., instructional conditions, as well as internal factors, i.e., metacognition and motivation, affect academic success and technology use (Winne, 1996; Gašević, Dawson, Rogers, & Gašević, 2016). These findings imply that failing to recognize differences between learners when designing learning analytics dashboards can result in inadequate tools that are quickly dismissed by learners.

Further shortcomings were identified by our literature review that analysed the use of educational concepts in existing LA dashboard (Jivet, Scheffel, Drachsler & Specht, 2017). The review was accepted as a full research paper to EC-TEL 2017 and a subsequent study analysing the extent to which educational concepts were considered into the evaluation of learning analytics dashboards was submitted at LAK 2018. Our main findings show that only half of the dashboards’ designs are grounded in learning sciences, but among the ones that are, *self-regulated learning* is the core theory that informs the design of LA dashboards. However, current designs offer some support only for the *performance* phase of the SRL cycle as defined by Zimmerman et al. (2000) (see Section 8.2.1.), as their main goal is to support awareness. Since very few dashboards are properly integrated into the learning environment or into the learning design, learners miss support in the other two phases, i.e. *forethought* and *self-reflection*. Furthermore, the majority of dashboards use comparison with peers as a representative frame of reference for evaluating their performance. Frames of references are anchor points that learners can use in order to make sense and evaluate the data displayed on the learning dashboard (Wise, 2014). Thus, there is a strong emphasis on comparison and competition with peers, although research in educational sciences identifies different sources of motivation for learners, i.e. performance and mastery achievement goal orientations (Elliot & McGregor, 2001).

Based on the previously identified problems of the field, we hypothesise that in order to provide well-balanced support for the development of SRL skills, a dashboard that simply displays metrics computed from learning data might not be sufficient. At the same time, we suspect that the strong focus on competition with peers as a motivational drive might have detrimental effects on the long term in learners as comparison to peers and “being better than others” would become the norm in terms of what defines a successful learner instead of enthusiasm for mastering knowledge, acquiring skills and developing competencies. Therefore, the goals of this PhD research project are to (i) develop and evaluate a comprehensive learning analytics dashboard that supports learners’ **self-regulated learning** and provides individualised feedback that takes into account learners' preferences for different **learning approaches,** e.g.,competitive learning or mastery learning, in order to provide alternative designs that do not focus only on competition with peers, and (ii) assess the effects of delivering individualised feedback via this dashboard on **learners' performance**. The selected target domain for the development of the dashboard is MOOCs (further explained in section 2.3). The main research question is the following:

> How do learning analytics dashboards need to be adapted in order to effectively support individual preferences for learners in self-regulated learning?

2 DESIGN AND METHODS

2.1 Theoretical foundation

Self-regulated learning has been a major research topic in learning sciences for more than 30 years with numerous theoretical models that have been published and empirically verified. Panadera's (2017) most recent work provides a comprehensive view on the six most popular models for SRL. Considering the learning paradigm behind each model and the comparison provided by Panadera (2017), for the development of our dashboard we will mainly rely on Zimmerman's model (2000). We have decided to use Zimmerman's model for the following reasons. Firstly, Zimmerman's model covers and balances the three main areas of SRL activity: cognition, motivation and emotion without emphasising one over the other. Secondly, motivation is well acknowledged in Zimmerman's model as it explicitly states the importance of goals and presents SRL as a goal-driven activity. Thirdly, Zimmerman's SRL is explained as a cyclical and interactionist process between cognitive processes, behaviour and environment, rather than a bi-directional flow like in the other models. Fourthly, the model has been extensively empirically validated, it's easy to understand and apply, resulting in the highest number of citations among the investigated models. Finally and most importantly, Panadera's (2017) review showed that there is a tendency for higher education students to have better results if they are exposed to interventions that are aiming at motivational and emotional aspects. Since our dashboard will be deployed in MOOCs where the majority of participants are adult learners (Macleod, Haywood, Woodgate & Alkhatnai, 2015), Zimmerman's model presents itself as the best fit. According to Zimmerman (2000), SRL is a social cognitive process that involves cognitive, emotional and behavioural processes. SRL is achieved in cycles of (i) forethought, i.e., learning task analysis and self-motivation beliefs, (ii) performance, i.e. execution of the learning task and progress monitoring, and (iii) self-reflection, i.e., self-judgement of outcomes and self-reaction.

At the same time, the design of the learning dashboard will be adapted based on learners' achievement goal orientation, a model developed by Elliot & McGregor (2001). This theory distinguishes between mastery and performance orientations as the motivation behind why one engages in an achievement task. In contrast to learners who set mastery goals and focus on learning the material and mastering knowledge, learners who have a performance orientation have a more competitive approach and are more focused on demonstrating their ability by measuring skill in comparison to others. We hypothesise that aligning a learner's achievement goal orientation with they way they are given feedback and the frames of reference used for self-evaluation can support learners to achieve their goals more effectively. From our literature review, we concluded that the design of current dashboards is more appealing to competitive learners, neglecting learners who have a preference for knowledge mastery.

2.2 Research questions

In order to support learners during SRL, our dashboard concept needs to effectively support each phase of the SRL cycle while considering the existence of different achievement goal orientations. Therefore, we have formulated the following research sub-questions to guide our research.

A. Forethought-F-RQ What data do learners need in a learning analytics dashboard in order to strategically plan their personal learning goals?

B. Performance-P-RQ Does visualisation tailored to the learner's learning approaches support the implementation of learners' goals?

C. Self-reflection-SR-RQ What are effective triggers within learning analytics dashboards for self-reflection that address learners' preferred learning approaches?

2.3 Approach and experimental design

The research project lasts four years and consists of a literature review and three prototype development iterations as illustrated in Figure 1. The goal of the literature review is to identify most common educational concepts used in the design and evaluation of learning dashboards and to pinpoint existing shortcomings in the design and use of learning dashboards in the educational practice. Initial results of the literature review were published at the EC-TEL 2017 conference (Jivet et al., 2017) and a subsequent study was submitted to LAK 2018.

In order to develop the envisioned learning analytics dashboard, we will follow an iterative design process with three successive design-evaluation cycles, following the design-based research methodology (Design-Based Research Collective, 2003) as shown in Figure 1. The effectiveness and usability of the implemented prototypes will be evaluated in randomised controlled trials deployed in MOOCs. Once the MOOCs have finished, learners will be asked to evaluate the dashboard prototypes using the evaluation framework for Learning Analytics[1] (EFLA) that also addresses SRL. In addition, the trace logs data collected will also be analysed as well as the final grades to determine differences in terms of engagement, performance and achievement between the treatment groups and the control group. Successful features will be integrated into the dashboard infrastructure and used in subsequent iterations. Each iteration will evaluate features belonging to one phase of the SRL cycle, answering the research questions outlined in the previous section. The first iteration targets the forethought phase (RQ-Forethought). The aim of these experiments is to identify learner data that are relevant and

[1] http://www.laceproject.eu/evaluation-framework-for-la/

valuable to learners based on their learning goals and dashboard features that support their goal setting. The MOOC learners will be divided into one control group (no feedback) and several treatment groups that will receive feedback according to their self-reported goal for the MOOC. The second iteration seeks to identify what kind of support learners need during the performance phase (RQ-Performance). By designing different versions of a dashboard that reflect learners' preferences between comparative and mastery attainment feedback and evaluating them in a real-life scenario, we aim to identify which form of data visualisation appeals most to individual learners for (self-)monitoring their progress and achievements. The final group of experiments concerns the self-reflection phase (RQ-Self-reflection). The purpose of these experiments is to identify effective triggers for reflection integrated into the dashboard design that take into account the learners' competitive and mastery orientations. We define effective triggers as dashboard features that drive learners to take action, change their learning strategy or do better planning.

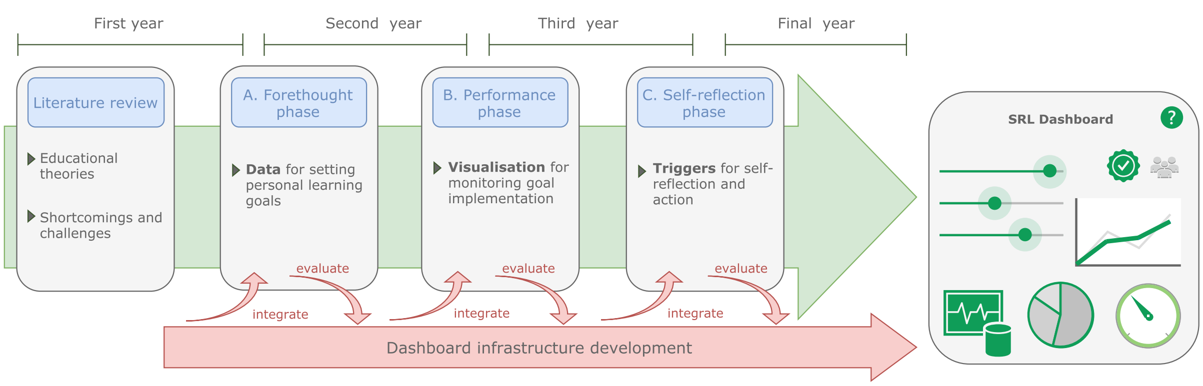


Figure 1. Timeline of the research project and the objectives tackled in each phase.

The selected target domain for the evaluation of the dashboard features are MOOCs. We selected MOOCs as one of the target domains for several reasons. Firstly, MOOCs present low completion rates due to low SRL skills of learners (Kizilcec, Perez-Sanagustin & Maldonado, 2017) and current MOOC platforms offer little support to develop learners' SRL skills. Secondly, MOOCs are usually attended by large and diverse populations of learners, contributing to the statistical relevance and validity of the experimental setups. Finally, due to their open nature, MOOCs are more flexible and present less restrictions when conducting randomised controlled trials, i.e. it is easier to add widgets during the run of the course or assigning learners to treatment groups, than courses provided as part of the curricula of institutions of higher education.

3 CONCLUSION

By developing a dashboard prototype that supports learners throughout each phase of the self-regulated learning cycle, we intend to offer a better integration of the dashboards within learners' learning activities, a feature currently lacking in the majority of learning dashboards. Furthermore, by tailoring the data and visual design of dashboards to the goals and motivations of learners we attempt to reduce the challenge of low uptake and usage that existing dashboards. In contrast to existing works, the starting point of this research is learning sciences. By considering the full self-regulated learning cycle, the final dashboard design aims to provide support for learners through each phase of the SRL cycle. As our literature review has revealed, the large majority of dashboards are designed to support awareness and to slightly trigger self-reflection, the final phase of the SRL cycle. At the same time, this work challenges the assumption that fostering competition through comparison with peers is a legitimate way of motivating all learners towards performance, a feature extremely common in the current design of learning dashboards. Throughout our research, we aim to design a learning analytics dashboard concept that supports learners who seek knowledge, skill and competency mastery, as well as learners that are motivated by competition with peers.

REFERENCES

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating effective learning analytics dashboards: Lessons learnt. In *Proc. of EC-TEL'16* (pp. 42-56). Springer.

Elliot, A. J., & McGregor, H. A. (2001). A 2×2 achievement goal framework. *Journal of personality and social psychology*, *80*(3), 501.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5-6), 304-317.

Gašević, D., Dawson, S., Rogers, T., & Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68-84.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64-71.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, *15*(3), 42.

Jansen, R.S., Van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Jnl. of Comp. in Higher Education*, *29*(1), 6-27.

Jivet, I., Scheffel, M., Drachsler, H., Specht, M. (2017). Awareness is not enough. Pitfalls of learning analytics dashboards in the educational practice. In *Proc. of EC-TEL 2017.* LNCS, Springer.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18-33.

Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, *47*, 90-97.

Macleod, H., Haywood, J., Woodgate, A., & Alkhatnai, M. (2015). Emerging patterns in MOOCs: Learners, course designs and directions. *TechTrends*, *59*(1), 56.

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, 8.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, *82*(1), 33.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., ... & Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*(1), 30-41.

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 5-8.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and individual differences*, *8*(4), 327-353.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of LAK'14* (pp. 203-211). ACM.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist*, *25*(1), 3-17.

Zimmerman, B. J. (2000). "Attaining self-regulation: a social cognitive perspective," in Handbook of Self-Regulation, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (San Diego, CA: Academic Press), 13–40.

Algorithmic Identity and Educational Practice

University of Newcastle
Janine.Arantes@uon.edu.au

ABSTRACT:

Big Data and algorithms can be used to personalize an educational environment and potentially improve the secondary school experience (ages 11-18). With individual learner profiles used to develop insights, the literature suggests that increased student engagement and the prediction of student behavior may be possible. Common to Massive Open Online Learning Courses (MOOCs) and other large learning platforms, personalization tools using learning analytics are at the cusp of mainstream implementation into secondary school educational practice. Similarly, whole school environments referred to as 'Startup Schools' are also being piloted in the United States. Beyond the educational context, learner profiles have been referred to as the 'algorithmic identity'. Described as the 'identity formed via algorithms that infer categories of identity on what would be otherwise anonymous data', this digital form of identity has also been theorized as a form of modulatory control and personalization tools described as the 'technological unconscious.' With human intervention minimized or completely removed through automated feedback loops, the 'technological unconscious' refers to algorithmic systems that shape the everyday in unseen and unknown ways. In the context of education, a lack of understanding may see teachers naively reproduce the same troubling dynamics that perpetuate a loss of human judgement, inequity and bias to perpetuate in the classroom, even in the absence of conscious prejudice. Exploring how teachers understand algorithms as part of their educational practice is therefore, both timely and significant. Guided by the foresight methodology 'Causal Layered Analysis' (CSA), the research will obtain transdisciplinary expert consensus on a future scenario in the context of education. Represented as a vignette, this scenario will be presented to Victorian secondary teachers during semi-structured in depth interviews, guided by Grounded Theory principles. Presented in the LAK2018 Doctoral Consortium as a poster, the research project aims to explore how Victorian secondary teachers understand algorithms as part of their educational practice when cognizant of broader ethical considerations.

Keywords: Algorithmic identity, algorithm, soft biopower, learning analytics, StartUp Schools, foresight, grounded theory

BACKGROUND

Alongside technological progression, the analysis of 'big data' has become part of every domain, including education. Far from just recording academic results and behaviors, Big Data, or data too large to be useful in standard databases (Power, 2014), is being used to personalize learning and predict student behavior (King, 2016). Operationalized via various mathematical rules called algorithms, personalization tools can provide learner insights through the creation of an 'algorithmic identity', without human intervention (Thompson, 2017). However, alongside the literature demonstrating significant educational benefits (Papamitsiou & Economides, 2014) (Siemens, 2013), there are numerous ethical concerns (Barocas and Selbst, 2016) (Beer, 2017) (Lightbourne, 2017).

The 'algorithmic identity' is defined as "an identity formation that works through mathematical algorithms to infer categories of identity on otherwise anonymous beings" (Cheney-Lippold, 2011, p. 160). Anonymous beings refer to a learner profile that is lacking a verified truth of identification, rather scored proxies are used to create the user profile (identity) (Zarsky, 2014). From a practical perspective, the algorithmic identity is assigned by the proprietor of the algorithm in tools such as learning analytics and educational data mining (Papamitsiou & Economides, 2014). From a theoretical perspective, multiple scholars argue that the algorithmic sorting of educational data to create algorithmic identities, sees a shift from disciplinary power, to power through modulatory control (Thompson, 2017) (Cheney-Lippold, 2011) (Lash, 2007). That is, the person can be controlled, not by discipline, punishment or praise, but by modulating the definition of the categories used to segment the data. For example, 'femaleness' could be represented by the proxy of 'engineering' or 'nursing', reducing the complexity of the term 'femaleness' to statistical correlations that may perpetuate stereotypes. This modulation of the categories creates what has been called 'soft biopower' (Cheney-Lippold, 2011), which is an indirect acceptance of directives and an unapparent, indirect form of control, where understanding or comprehension of how decisions are made is unknown.

Problem Statement

Currently, algorithmic approaches to secondary education are being implemented into Australian classrooms. Theorised as a form of modulatory control, the literature suggests that society is largely unaware of associated concerns, although increasingly impacted. Concomitant expert concern as applied to the educational context has not transmuted from the theoretical to the evidencing of teachers' practical understanding.

Goals

- To consolidate transdisciplinary opinion related to algorithmic identity and soft biopower in the context of secondary education (11-18 year old students)
- To provide foresight into how identified concerns may transmute into secondary educational practice
- To bring about debate and discussion of how learning analytics are understood in the context of secondary education
- To challenge assumptions or concerns held with learning analytics in the context of secondary educational practice

Research Question

How do teachers understand algorithms as part of their educational practice?

Definition of Terms

- *Algorithmic Identity:* Algorithmic Identity is "an identity formation that works through mathematical algorithms to infer categories of identity on otherwise anonymous beings" (Cheney-Lippold, 2011, p. 160)
- *Soft bio-power:* Soft bio power is where there is an indirect acceptance of directives issued by algorithmic authoritarian bodies, forming an unapparent and indirect form of control. (Cheney-Lippold, 2011)
- *Startup Schools:* Startup Schools see entrepreneurs create an educational environment where data is continuously collected about how students learn, their preferences and general aptitude through pattern analysis and predictive techniques (Williamson, 2016) (Williamson, 2018). An example is *AltSchool* (https://www.altschool.com/)

CURRENT KNOWLEDGE

The literature provides multiple concrete empirical examples explaining tools used to improve educational practice. Papamitsiou and Economides (2014) segment personalization tools into two complimentary areas: Learning Analytics (LA) and Educational Data Mining (EDM). Jeong & Biswas (2008) demonstrate that higher learning strategies can be inferred, and retained by students after feedback prompts are removed. Blikstein (2011) shows that teachers can understand students' trajectories and infer strategies through automatically generated data, and Kizilcec, Piech, & Schneider (2013) provide a framework for student engagement based on MOOC datasets. Abdous, Wu, & Yen (2012) explore performance prediction, to reveal a greater understanding of students' learning behaviors through a combination of educational data mining and learning analytics, and Romero-Zaldivar, Pardo, Burgos, & Kloos (2012) determine how the mining and analysis of student data can make predictions of student academic achievement. West, Heath, & Huijser (2016) explore retention prediction to introduce a dialogical tool that claims to advance student retention and Giesbers, Rienties, Tempelaar, & Gijselaers (2013) explore how engagement in learning activities aligns with retention and metacognition has been investigated. Clow (2013) introduces the concept of a 'funnel of participation' to make evident the usefulness and motivational capabilities various aspects of Massive Online Open Courses (MOOC) and Merceron & Yacef (2008) demonstrate how teachers can be informed about the use of extra learning material in Learning Management Systems (LMS) such as *Moodle*. Clow (2013) compiles various types of tools to qualify the appropriateness of feedback and Blikstein (2011) examines adaptive testing in formative assessment. Danaher et al. (2017) highlights that 'techno-utopianism' is prevalent in education, seeing technology as a solution to education's problems, although potentially obscuring ethical considerations.

Beyond the educational context, the algorithmic identity has been discussed in relation to power from multiple theoretical viewpoints (Lash, 2007) (Cheney-Lippold, 2011) (Thompson, 2017), leading to discourse related to ethical considerations. Barocas & Selbst (2016) argue that algorithmic categorizations deepen divides and increases disadvantage, Bucher (2012) suggests that there is an 'architectural structuring of visibility' with high socio-economic groups exposed to digitally different opportunities than low, and Beer (2017) argues that males see a different internet to females. From an educational perspective, ethical concerns include business models disguised in educational products (Gutstein, 2012) (Lindh & Nolin, 2016) and the ethics of surveillance (Williamson, 2017a) and the enablement of control (Thompson, 2017). Ruppert et al. (2015) suggest that the connectedness and interdependency of Big Data, has resulted in the capacity of these tools to establish social relationships. As a result, various computational techniques have become not only a means to manage students, but also a source of social control (Williamson, 2017b). Some argue that we may see a 'degovernmentalization' of education governance and policy (Olmedo, 2014), where there is "ultimately…a 'free', but constantly conditioned, user" (Cheney-Lippold, 2011, p. 178). Supporting this notion is Margetts & Dunleavy (2013), who suggest that personalization technologies enable a form of algorithmic governance. Furthermore Pasquale (2015) and Beer (2017) refer to the notion of the 'technological unconscious', where technologies that shape the everyday can act in unseen and unknown ways and Danaher et al. (2017) highlight that where there is surveillance without human participation, nor comprehension, the legitimacy of the power enabled should be questioned.

Existing solutions

Internationally, the OECD has released a socio-emotional skills assessment framework and the UK government's Behavioral Insights team or 'Nudge Unit' has been making use of publically available data to better help policy makers and educational practitioners (Williamson, 2016). The Institute of Electrical and Electronics Engineers (IEEE) are developing ethical design standards to allow organizations to be certified according to various ethical standards (IEEE, 2017). Australian legislation and policy does not contain explicit standards related to algorithms in education, and Victorian education solutions, refer back to state and federal solutions and individual school policy. The Victorian secondary school context was chosen due to the operational consideration of completing face-to-face interviews and the researcher's knowledge of the Victorian secondary school system.

Legislative and Rights based solutions

Yeung (2017) discusses how the embedding of algorithms in adver-games, allows children to be 'hyper-nudged' and as such alter their behavior in a predictable manner. These influences have been seen by some as a growing and pervasive commercialization of childhood and a violation to the child's right to identity (Livingstone, Mascheroni, & Staksrud, 2017). Brownsword (2009) suggests that the algorithmic identity "inhibits non-conformist life choices; and…generates inaccurate profiles that result in decisions that impinge on both universal and local rights" (p. 248). Internationally, The European Union (EU) General Data Protection Regulations (GDPR), effective May 25 2018, aims to provide data security to residents of the European Union. Within this framework, EU residents are provided with various rights, such as the power to challenge decisions informed by algorithmic insights (Goodman & Flaxman, 2016). Australian's currently have the right to access and edit their data, but not the right to challenge automated decisions or stop a company collecting data about them (Productivity Commission, 2017). A contingency has been suggested by the Australian Law Reform Commission that the Privacy Act be amended to incorporate a capacity test, to assess individuals under the age of 18 as part of Australian Uniform Evidence Law (ALRC, 2017), thus moving towards more child-specific privacy protection when online (Mercantile, 2017).

Design solutions

(Hallinan & Striphas, 2014) suggest that there is a need to improve conceptual and semantic work to render algorithmic systems legible as forms of cultural decision making. (O'Neil, 2016) argues that the algorithmic systems "pick(ed) proxies that seem(ed) to correlate with success" and that those who control the code can more readily manipulate the complex realties they represent. Therefore with the algorithmic identity removing the singular truth of identification, and replacing it with a representation of mass aggregated data sets, changes in design have offered a solution. Discrimination aware data mining (Berendt, 2014) and Privacy by Design principles (Rubenstein & Good, 2013) claim to give control to the individual, so that they can understand and challenge a decision made by a proprietorially protected algorithm (Rubenstein & Good, 2013). The Institute of Electrical and Electronics Engineers (IEEE) are developing ethical design standards, which will allow organizations to demonstrate that they conform to various ethical standards. For example, the Standard for Child and Student Data Governance (IEEE P7004) focusses on transparency in student and child data collection and analysis, and the Standard for Personal Data Artificial Intelligence (AI) Agent focusses on maintaining human judgement in all decision making.

A UNIQUE APPROACH

Personalized learning enables algorithmic governance systems which incorporate surveillance as a ubiquitous entity in the everyday (Willson (2014), but are not part of the day to day consciousness (Hayles, 2006). This highlights a need to explore how teachers understand algorithms (used in learning analytics) as part of their educational practice, so as to be able to critically construct debate that can keep pace with the advances in technology. Current solutions in Australia include the incorporation of preventative design methods, yet legislation specific to offer a transparent ability to scrutinize, understand or challenge decisions as seen in the EU GDPR does not exist. However, where law and design can protect users, should the legislation or design be flawed and human intervention (such as in feedback loops) or comprehension is lacking, those impacted cannot understand decisions and as such debate concerns or have informed choice. This can lead to severe repercussions. For example, research by Lightbourne (2017) argues that an algorithmic assessment showed a greater threat of recidivist risk for a defendant, due to the color of his skin and that when the judiciary were informed of this potential, their 'understanding' that data is objective influenced the ultimate decision. This may be a valid concern in the context of secondary education, and this research project aims to address this problem in a unique manner. Unique in that, the findings don't seek to inform policy or legislation, rather concomitant expert opinion will provide foresight into a possible, preferable and probable scenario to guide teacher understanding and knowledge development for further debate. By involving teachers in the construction of their understanding from a critical perspective, the research aims to empower teachers at the vanguard to robustly discuss, debate and gain further knowledge, so that they can challenge their assumptions and understanding of algorithms, Big Data and associated ethical considerations.

Research Methodology

Based on a foresight framework, called *Causal Layered Analysis (CSA)* (Inayatullah, 1998), the research aims to provide a systematic and structured analysis of expert opinion of current and future trends to develop a possible, preferable and probable scenario to present to secondary teachers for further analysis (Refer to Figure 2. 'Research Plan'). This design was chosen to overcome challenges with researching algorithms from a critical perspective, such as transparency, interpreting code and access to how they were formed. The methodological approach focusses on unpacking the socio-technical assemblage to study real world effects by exploring the social impacts of algorithms within specific contexts (Kitchin, 2017). However, as the results to be understood are not obvious, the methodology aims to establish a factual correlation via foresight methodology initially (Slaughter, 1997). This will characterize the algorithmic identity and soft bio power in the educational context prior to analysis within the specific context of education. The main data collection will occur during the interviews, guided by Grounded Theory principles (Strauss & Corbin, 1994). The research situation will be shaped via the introduction of a vignette developed from the Delphi findings (Okoli & Pawlowski, 2004). A brief summary of details are included in Figure 1. "Methodology".

Figure 1. Methodology

	Delphi	**Semi Structured In Depth Interviews**
Participants	Trans-disciplinary experts with an understanding of the algorithmic identity	Victorian secondary teachers (teachers represent: teachers, senior teachers, principals, deputy principals, educational support staff)
Procedure	6 (max) x 15 minute	Face to face interviews of approx. 2 hours, including

	questionnaires via hyperlink over 3 months	open and closed ended questions, Likert-type scales, mediated by a vignette based on the Delphi findings Estimated high impact, low probability event discussion (QUAL quan). Grounded Theory (Strauss & Corbin, 1994), with theoretical saturation of 25-30 (TBC)
Analysis	Guided by (Okoli & Pawlowski, 2004) Two by two matrix, cross impact analysis	Grounded Theory, ongoing analysis and coding (open, axial, selective). An iterative process, aiming to develop a substantive theory that will be presented in the form of statements of importance.
Ethical issues	Corporate IP, anonymity, voluntary participation.	Organizational consent to participate, storage of personal information, UoN ethics, understanding of procedures and potential risk. Mitigation of risk.
Limitations	Limited by current knowledge. Potential for a manipulated consensus.	Teacher understanding is a multifaceted cognitive and behavioral construct, thus various features will need to be defined to create the framing and increase validity (e.g. the vignette).

Ethnographic design, focus groups and surveys were also considered. Semi structured in depth interviews, guided by Grounded Theory was chosen to provide structure and increase methodological rigor. These methods also align with the underpinning epistemology, allowing the construction of knowledge from a critical perspective. This more discursive approach was also chosen to explore the social impacts of algorithms within the specific context of education as per Kitchin's (2017) guidance and increase internal validity through the use of a vignette. Understanding that the transferability of the emerging themes is limited and only bare relevance to those teachers directly involved, the statements of importance are anticipated to be used as a resource for further debate and wider discourse.

Current status of work

Figure 2. Research Plan (and current status of work)

CSA LEVEL	**Problem**	**Provider**	**Solution**	**Participants**	**Method**	**Analysis**	**Progress**
Litany	The visible problem appears too challenging to solve.	Media / Corporate	Short term approach	Non Scholarly Literature	Scope		2017
Social Causes	The problem is presented as a systemic or procedur	Academic / Government	Systematic solutions	Scholarly peer reviewed literature	Literature Review		2017

	al issue						
Worldview	The problem is constituted by the framing of the analysis (algo ID + soft biopower)	Trans disciplinary experts including beyond the dominant language	Exploring social impacts of algorithms Consideration of alternate views	10-18 Transdisciplinary experts with an understanding of algorithmic identity	Purposeful sampling Delphi Forecasting (web based)		2018
Metaphor / Myth	The problem is constituted by a core story (vignette)	The teachers' collective unconscious	Considering social impacts of algorithms in specific contexts Alternative stories / futures	30 (estimated) teachers* from the Victorian Secondary Education system *including principals, deputies, support staff	Snowball Sampling Semi Structured In depth Interviews (Predominantly Qualitative)	Grounded Theory	2019

REFERENCES

Abdous, M., Wu, H., & Yen, C. (2012). Using data mining for predicting relationships between online question theme and final grade. *Journal of Educational Technology & Society, 15*(3), 77-88. doi:10.1.1.298.8417

Barocas, S. & Selbst, A. (2016). Big Data's Disparate Impact. *California Law Review, 104*(3), 671-732. doi:10.15779/Z38BG31

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society, 11*(6), 985-1002. doi:10.1177/1461444809336551

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society, 20*(1), 1-13. doi:10.1080/1369118X.2016.1216147

Berendt, B. & Preibusch, S. (2014). Better decision support through exploratory discrimination-aware data mining: foundations and empirical evidence. *Artificial Intelligence Law, 22*, 175-209. doi:10.1007/s10506-013-9152-0

Blikstein, P. (2011). *Using learning analytics to assess students' behavior in open-ended programming tasks.* Paper presented at the Proceedings of the 1st international conference on learning analytics and knowledge.

Bostrom, N. (2014). *Superintelligence. Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Brownsword, R. (2009). Friends, Romans, Countrymen: Is there a Universal Right to Identity? *Law, Innovation & Technology., 1*(2), 223-249. doi:10.1080/17579961.2009.11428372

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society, 14*(7), 1164-1180. doi: 10.1177/1461444812440159

Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory, Culture & Society, 28*(6), 164-181. doi:10.1177/0263276411424420

Clow, D. (2013, April). *MOOCs and the funnel of participation.* Paper presented at the Third International Conference on Learning Analytics and Knowledge (2013), Leuven, Belgium.

Danaher, J., Hogan, M. J., Noone, C., Kennedy, R., Behan, A., De Paor, A., Felzmann, H., Haklay, M., Khoo, S-M., Morison, J., Murphy, M.H., O'Brolchain, N. & Shankar, K. (2017). Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big Data & Society, 4*(2), 1-21. doi:10.1177/2053951717726554

Freire, P. (2000). *Pedagogy of the oppressed*. New York, USA: The Continuum International Publishing Group Inc.

Giesbers, B., Rienties, B., Tempelaar, D., & Gijselaers, W. (2013). Investigating the relations between motivation, tool use, participation, and performance in an e-learning course using web-videoconferencing. *Computers in Human Behavior, 29*(1), 285-292. doi:10.1016/j.chb.2012.09.005

Gutstein, D. (2012). *Pearson's plan to control education. Report to the B.C. Teachers' Federation.* British Columbia: British Columbia Teachers' Federation.

Hallinan, B. & Striphas, T. (2014). Recommended for you: The Netflix Prize and the production of algorithmic culture. *New Media & Society, 18*(1), 117-137. doi:10.1177/1461444814538646

Hayles, N. K. (2006). Unfinished work from cyborg to cognisphere. *Theory, Culture & Society, 23*(7-8), 159-166. doi:10.1177/0263276406069229

IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2. IEEE, 2017. http://standards. ieee.org/develop/indconn/ec/autonomous_ systems.html

Inayatullah, S. (1998). Causal layered analysis: Poststructuralism as method. *Futures, 30*(8), 815-829. doi:10.1016/S0016-3287(98)00086-X

Jeong, H. & Biswas, G. (2008). *Mining student behavior models in learning-by-teaching environments.* Paper presented at the Educational Data Mining 2008.

King, M., Cave, R., Foden M. & Stent, M. (2016). *Personalised education. From curriculum to career with cognitive systems* (IBM Ed.). Portsmouth, United Kingdom: IBM Corporation.

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society, 20*(1), 14-29. doi:10.1080/1369118X.2016.1154087

Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). *Deconstructing disengagement: analyzing learner subpopulations in massive open online courses.* Paper presented at the Proceedings of the third international conference on learning analytics and knowledge, Leuven, Belgium.

Lash, S. (2007). Power after Hegemony : Cultural Studies in Mutation? *Theory Culture Society, 24*(55), 1-25. doi:10.1177/0263276407075956

Lindh, M. & Nolin, J. (2016). Information We Collect: Surveillance and Privacy in the Implementation of Google Apps for Education. *European Educational Research Journal, 15*(6), 644-663. doi:10.1177/1474904116654917

Livingstone, S., Mascheroni, G., & Staksrud, E. (2017). European research on children's internet use: Assessing the past and anticipating the future. *New Media and Society*, 1-20. doi:10.1177/1461444816685930

Margetts, H., & Dunleavy, P. (2013). The second wave of digital-era governance: a quasi-paradigm for government on the Web. *Philosophical Transactions of the Royal Society Association, 371*(1987). doi:10.1098/rsta.2012.0382

Mercantile, M. (2017). From universal towards child-specific protection of the right to privacy online: Dilemmas in the EU General Data Protection Regulation. *New Media and Society*.

Merceron, A., & Yacef, K. (2008, January). *Interestingness measures for association rules in educational data.* Paper presented at the The 1st International Conference on Educational Data Mining, Montréal, Québec, Canada.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Broadway Books.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & management, 42*(1), 15-29. doi:10.1016/j.im.2003.11.002

Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society, 17*(4), 49.

Productivity Commission. (2017). *Data Availability and Use, Report No. 82* (AGPC Ed.). Canberra: Australian Goverment Productivity Commission.

Power, D. J. (2014). Using ‘Big Data’ for analytics and decision support. *Journal of Decision Systems, 23*(2), 222-228. doi:10.1080/12460125.2014.888848

Romero-Zaldivar, V.-A., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education, 58*(4), 1058-1067. doi:10.1016/j.compedu.2011.12.003

Rubenstein, I., & Good, N. (2013). Privacy by Design: A Counterfactual Analysis of Google and Facebook Privacy Incidents. *Berkeley Law Journal, 28*(2 Fall), 1-83. doi:10.15779/Z38G11N

Ruppert, E., Harvey, P., Lury, C., Mackenzie, A., McNally, R., Baker, S. A., Kallianos, Y. & Lewis, C. (2015). *A Social Framework for Big Data*. warwick.ac.uk: Warwick Centre for Interdisciplinary Methodologies.

Selwyn, N. (2016). *Is technology good for education?* Cambridge: Polity Press.

Siemens, G. (2013). Learning Analytics. *American Behavioral Scientist, 57*(10), 1380-1400. doi:10.1177/0002764213498851

Slaughter, R. A. (1997). Developing and applying strategic foresight. *ABN Report, 5*(10), 13-27.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of qualitative research, 17*, 273-285.

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International journal of educational research, 27*(4), 293-302.

Thompson, G. (2017). Computer adaptive testing, big data and algorithmic approaches to education. *British Journal of Sociology of Education, 38*(6), 827-840. doi:10.1080/01425692.2016.1158640

West, D., Heath, D., & Huijser, H. (2016). Let's Talk Learning Analytics: A Framework for Implementation in Relation to Student Retention. *Online Learning, 20*(2), 150-170. doi:10.1145/2460296.2460299

Williamson, B. (2016). Coding the biodigital child: the biopolitics and pedagogic strategies of educational data science. *Pedagogy, Culture & Society, 24*(3), 401-416. doi:10.1080/14681366.2016.1175499

Williamson, B. (2017a). *Big Data in Education: The digital future of learning, policy and practice*. London: SAGE.

Williamson, B. (2017b). Moulding student emotions through computational psychology: affective learning technologies and algorithmic governance. *Educational Media International, 54*(4), 267-288. doi:10.1080/09523987.2017.1407080

Willson, M. (2014). The politics of social filtering. *Convergence-the International Journal of Research into New Media Technologies, 20*(2), 218-232. doi:10.1177/1354856513479761

Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society, 20*(1), 118-136.

Zarsky, T. (2014). Understanding Discrimination in the Scored Society. *Washington Law Review, 89*(4). doi:https://ssrn.com/abstract=2550248

Understanding Collaborative Learning: A Learning Analytics Approach

Author: Ed Fincham
University of Edinburgh
Ed.Fincham@ed.ac.uk

ABSTRACT: The development of educational technology has provided ample opportunity for the deployment of collaborative learning activities and has attracted the attention of researchers from fields ranging from computer supported collaborative learning, to psychology and computer science. Although it has produced many advancements in research and practice, existing research has left many pertinent questions regarding collaborative learning unanswered. In particular, there is little understanding of how individual self-regulated learning relates to the collaborative construction of knowledge within groups, how roles emerge within these groups, and how group regulation relates to individual self-regulation. To answer these questions, we propose to augment existing methods from social network analysis with behavioural information to investigate how collective regulation arises from the actions and self-regulating behaviour of individual agents. In doing so, we seek to identify a mechanism by which collaboration and group regulation occurs.

Keywords: Collaborative learning, self-regulated learning, social network analysis, learning strategies, discussion forum

1 INTRODUCTION

From education to professional work, socially enriched environments have become increasingly pervasive of contemporary life. Accordingly, collaborative learning has attracted much attention and numerous studies have documented a wide range of potential benefits. Following Griffin & Care (2015), collaboration encompasses two separate but interrelated domains: the social, and the cognitive. While the former relates to individual participation, the ability to take others' perspective, and social regulation, the latter relates to self-regulation of learning, the ability to solve problems, and the construction of knowledge. As isolated topics, there is extensive literature on both. However, there has been limited research into the interaction of these two domains (Hadwin & Oshige, 2011). To address this limitation, we aim to augment existing methods from social network analysis (SNA) with behavioural information to investigate how collective regulation arises from the actions and self-regulating behaviour of individual agents. In doing so, we seek to identify a mechanism by which collaboration and group regulation occurs. To this end, our investigation will rely upon data drawn from online learning platforms such as MOOCs. Of particular interest are the interactions that students have with the platform and the interactions students have with each other whilst using the platform.

2 THEORETICAL FRAMEWORK

Historically, the educational research literature conceptualised self-regulated learning (SRL) as a cognitive-constructive activity performed at the level of the individual (Winne, 1997). Contemporary perspectives, however, consider social influences to be at the core of SRL (Hadwin & Oshige, 2011). One such framework that emphasises the collaborative construction of knowledge is socially shared regulation of learning (SSRL). In essence, SSRL is collective regulation where regulatory processes are constructed collaboratively and the products – cognition – are shared. However, this is not to displace individual self-regulation, but rather to re-frame it within the context of social goals and a shared awareness of those goals (Hadwin & Oshige, 2011).

Research into SSRL often occurs in technology-based learning environments, where social exchanges, co-construction of knowledge, and student behaviour can be more easily traced (Hadwin & Oshige, 2011). Often such exchanges occur within computer-supported collaborative learning (CSCL) settings. CSCL offers students a variety of tools for constructing, representing, and mediating knowledge between peers in a collaborative environment. Accordingly, CSCL environments represent a shift from the traditional roles of teachers and learners in classroom instruction, towards partially autonomous knowledge building communities (Strijbos & De Laat, 2010). Considering the research into CSCL from a SSRL perspective reveals a substantive gap in the literature. Within SSRL, the focus is on collective regulation and there is little theoretical consideration of the impact of roles within the community. The concept of roles has had a major influence on our understanding of how individuals interact and their functions within a community (Biddle, 1986; Strijbos & De Laat, 2010).

Existing methods to identify roles rely on methods drawn from SNA but typically remain at a surface level of description, such as whether or not individuals are central and whether or not they frequently participate (Strijbos & Weinberger, 2010). Accordingly, more advanced techniques are required to help us identify which facets of individual behaviour are most predictive of an individual's role and how roles impact on the learning and self-regulating behaviour of other individuals within a group. To achieve this, our research will rely on methods from SNA, but will incorporate external metrics of learner behaviour, namely, social participation and learning strategies.

2.1 Social Participation Strategies

In the identification of roles, the existing literature places an almost exclusive emphasis on actions; what people do, with whom they do it. But this overlooks a vital aspect of participation, which may be characterised as "online listening behaviours" (Wise et al., 2013): how learners interact with the existing discussion and build knowledge, but without directly engaging with others. The greater control that students exert over their decision-making in this realm makes it an important facet of participation in online discussion forums and collaborative settings (Wise et al., 2013).

2.2 Learning Strategies

Collaborative learning within an SSRL setting encompasses both collective and individual regulation. Thus, to understand how group processes relate to individuals, a metric of individual behaviour that captures their self-regulating activity is required. This metric is learning strategies. While there are an abundance of definitions of learning strategies, the broad account provided by Weinstein et al. (2012, p. 227) whereby a learning strategy represents "any thoughts, behaviours, beliefs or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills" is often adopted.

Existing research into the identification of learning strategies in online learning environments makes frequent use of trace data (Winne, 2013). This research has amply demonstrated the ability of such techniques to extract latent representations (Jeong et al., 2008; Jovanović et al., 2017), there are numerous shortcomings. Firstly, while the existing literature has broadly categorised learning strategies into three types (Kovanović et al., [Submitted]), there are a multitude which vary on an inter- and intra-individual basis. Secondly, these representations are often static, extracted from arbitrary time periods. Finally, extracted from the particular actions and resources of a given course design or online setting, representations are difficult to generalise and compare across learning environments.

A more nuanced analytical framework requires a deeper theoretical understanding of the latent construct we are trying to extract. Rather than focusing on analytical methods to segment patterns of behaviour into groups or clusters, we may consider the depth of cognitive processing or overall sophistication a learning strategy encompasses. Such a construct may be described as a complex constellation of attitudes, knowledge and motivation that determine the degree to which a student cognitively engages with learning resources. Following Milligan and Griffin (2016), a plausible candidate for developing a broader analytical understanding of learning strategies is Item Response Theory (IRT) (Messick 1995). Using this approach, an individual's "item responses", or actions on a set of given indicators, are considered to be observable manifestations of a latent construct, here the sophistication of their learning strategy.

The application of Milligan and Griffin's (2016) methodology to the identification and analysis of learning strategies would address all three of the identified shortcomings with existing methods. Firstly, strategies would be identified over a continuum rather than, say, as a one of K cluster assignment. Secondly, the analysis could be run over numerous, variable time-frames to provide a more dynamic understanding of student learning strategies for instructors and the planning of interventions. Finally, as the results of Milligan and Griffin (2016) indicate, there is scope for generalisation between different courses and cohorts.

2.3 Social Network Analysis

While the insights gleaned from social participation and learning strategies are limited to the individual, SNA can combine this information with that of how individuals interact. This enables us to extend analyses beyond the individual and focus instead on how groups arise. This is especially so in blended and online learning environments where SNA can aid our understanding of how individual roles emerge and the impact such roles have on the collaborative efforts of a group.

Within SNA, mathematical models are often used to identify and describe relationships. However, the problem with such methods is that they are purely descriptive and offer no insight into the evolution of a given network, nor the propensity for groups to exhibit certain types of collaborative behaviour, nor which facets of individual behaviour are most predictive of an individual's role in a group. Instead, to understand these phenomena we must rely on statistical models (Goodreau et al., 2009). One commonly proposed method which could explain the formation of the observed network are Exponential Random Graph models (ERGMs).

Introduced by Frank and Strauss (1986), ERGMs belong to a family of probability models that allow for generalisable inferences over the structural foundations of social behavioural patterns within networks. ERGMs treat network ties as random variables, and model the overall network structure through a set of local network processes. The model assumes each tie within these processes is conditionally dependent, indicating that "empirical network ties do not form at random, but they self-organise into various patterns arising from underlying social processes" (Wang et al. 2013, p. 3).

3 RESEARCH QUESTIONS

While there has been much productive research into collaborative learning, there are a number of pertinent questions that remain unanswered. In particular, understanding how self-regulation occurs within collaborative learning environments, and how roles and group regulation relates to individual self-regulation of learning strategies. As our metric of self-regulated change over time (Winne 2013), developing a theoretically grounded framework for the analysis and interpretation of learning strategies is an important first step:

Research Question 1: *Can a generalised metric of Learning Strategies that allows for the assessment of an individual's learning strategy choices across a range of diverse courses be developed? Can this metric be implemented across a range of time frames to provide a more dynamic understanding of how an individual's learning strategies evolve over time?*

With a more nuanced understanding of self-regulation, we can then begin to investigate how individual self-regulation relates to group regulation. By combining statistical models of network formation with individuals' strategies relating to social participation and learning we can also broaden our understanding of what roles are and how they impact on others within a group. This prompts our second research question:

Research Question 2: *Can we incorporate user information that is external to the network structure, such as social participation and learning strategies, to further our understanding of how roles emerge? Furthermore, are collaborative learning environments affected by such external information? For instance, do network exhibit selective mixing regarding learning strategies?*

By developing a more generalisable framework for individual self-regulation and developing the analysis of roles within groups, we may be able to better understand collective regulation and how the actions of individuals affects change in other members of the group:

Research Question 3: *Through understanding individual self-regulation and the impact of roles within groups, can we develop a deeper understanding of SSRL? By analysing group dynamics over time can we identify how individual self-regulation interacts with group dynamics to foster the collaborative construction of knowledge and socially shared regulation?*

A central methodological development that our first research question proposes is the reformulation of the contemporary analytical framework regarding learning strategies. Rather than rely on, say, HMMs and clustering, we wish to identify and assess learning strategies using Item Response Theory (IRT) (Milligan & Griffin, 2016). Under this framework, discrete "item responses" on a given set of indicators are taken to be observable manifestations of a latent construct or ability. It may be possible to use IRT to define a theoretical framework that, given a suitable set of indicators, relates an individual's actions to the overall sophistication of their learning strategy. A key part of this research will be investigating the extent to which such indicators can be automatically identified. Such an approach has the benefits of allowing learning strategies to be assessed over variable time frames and, by proposing a common metric, generalisable across courses and cohorts.

Addressing our second research question, we will expand upon the existing literature regarding the identification of roles by incorporating behavioural information, as mediated through social participation strategies and learning strategies. The former of which may be extracted using the data mining methods outlined in the preceding paragraph. Using models such as ERGMs, we will investigate the statistical likelihood of network formation processes and assess the extent to which external metrics of individual behaviour and engagement are predictive of their role within groups.

Addressing our third research question will draw on our research into both learning strategies and emergent roles. This work will seek to investigate how collective, collaborative regulation of a group emerges from the self-regulating activity of individuals. Within a network, we will investigate how individuals' self-regulating activity, as measured by their changing learning strategies, changes with network formation, and to what extent individuals' roles, engagement strategies, and self-regulating activity are predictive of the self-regulating activity of others. Network clustering techniques may also be used to identify the presence and extent of group formation within a network, and may provide some insight into how these groups are formed.

4 PRELIMINARY RESULTS

Our first research question was motivated by the theoretical limitations we faced regarding the interpretation of learning strategies in a study assessing the effect of interventions on learning strategies (Fincham et al., [Submitted_1]). Our study looked at the effects of a formative feedback intervention on the learning strategies that students adopted in a flipped classroom setting. While we found that framing learning strategies as dynamic constructs offered insight into how student's self-regulating behaviour changed following an intervention, our results highlighted the necessity for a richer theoretical framework that can be applied over a variety of time scales and is capable of providing cross-course comparisons. Addressing this limitation is a central part of our research agenda.

Our second study (Fincham et al., [Submitted_2]) acted as an exploratory precursor to our broader research agenda. We investigated the impact that a broad range of methods for extracting social networks had on the resultant structural and statistical properties of social networks and the association between measures of centrality and academic performance. The study was conducted on data from two distinct learning environments and found that social tie definition plays an important role in shaping the results of such analyses, to the extent that the association between centrality and academic performance can in some cases be reversed. These findings will inform the methodological decisions made in pursuit of our second and third research questions.

REFERENCES

Biddle B. J. (1986). Recent development in role theory. *Annual Review of Sociology. 12,* 67–92. http://dx.doi.org/10.1146/annurev.so.12.080186.000435

Fincham E., Gašević D., Jovanović J., & Pardo A. ([Submitted_1]). From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations.

Fincham E., Gašević D., & Pardo A. ([Submitted_2]). From Social Ties to Network Processes: Do Tie Definitions Matter?

Frank, O. D., & Strauss, D. (1986). Markov Graphs. *Journal of the American Statistical* Association. *81*(395). 832–842. http://dx.doi.org/10.1080/01621459.1986.10478342

Goodreau S., Kitts J., & Morris M. (2009) Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks. *Demography. 46* (1), 103–125. http://dx.doi.org/10.1353/dem.0.0045

Griffin P., & Care E. (2015). Assessment and Teaching of 21st Century Skills: Methods and Approaches. Springer.

Hadwin A., & Oshige M. (2011). Self-regulation, co-regulation, and socially-shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record. 113* (2), 240–264

Jeong H., Gupta A., Roscoe R., Wagster J., Biswas G., & Schwartz D. (2008). Using Hidden Markov Models to Characterize Student Behaviors in Learning-by-Teaching Environments. Springer Berlin Heidelberg, 614–625

Jovanović J., Gašević D., Dawson S., Pardo A., & Mirriahi N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education. 33,* 74–85. http://dx.doi.org/10.1016/j.iheduc.2017.02.001

Kizilcec R. F., Pérez-Sanagustín M., & Maldonado J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education. 104*, 18–33. http://dx.doi.org/10.1016/j.compedu.2016.10.001

Kovanović V., Joksimović S., Poquet O., Hennis T., de Vries P., Hatala M., Dawson S., Siemens G., & Gašević D. ([Submitted]). The role of technology use on shaping student learning experience in MOOCs.

Messick S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*. 5–8. http://dx.doi.org/10.1111/j.1745-3992.1995.tb00881.x

Milligan S., & Griffin P. (2016). Understanding learning and learning design in moocs: A measurement-based interpretation. *Journal of Learning Analytics. 3* (2).

Strijbos J. W., & De Laat M. F. (2010). Developing the role concept for computer-supported collaborative learning: An explorative synthesis. *Computers in Human Behavior. 26*, 495–505

Strijbos J. W., & Weinberger A. (2010). Emerging and scripted roles in computer-supported collaborative learning. *Computers in Human Behaviour. 26*, 491–494

Wang P., Robins G., Pattison P., & Lazega E. (2013). Exponential random graph models for multilevel networks. *Social Networks. 35* (1), 96–115. http://dx.doi.org/10.1016/j.socnet.2013.01.004

Weinstein C. E., Husman J., & Dierking D. (2012). Self-Regulation Interventions with a Focus on Learning Strategies. Handbook of Self-regulation. 727-747

Winne P. H. (1997). Experimenting to bootstrap self-regulated learning. *J. Educ. Psychol. 89,* 79–410

Winne P. H. (2013). Learning Strategies, Study Skills, and Self-Regulated Learning in Postsecondary Education, Springer Netherlands, Dordrecht, 377–403

Wise A. F., Speer J., Marbouti F., & Hsiao Y.-T. (2013). Broadening the notion of participation in online discussions: examining patterns in learners online listening behaviors. *Instructional Science. 41*, 323–343. http://dx.doi.org/10.1007/s11251-012-9230-9

Enhancing Student Success in Open Online Education Based on Theories of Self-Regulation, Motivation and Cognitive Load: A Learning Analytics Approach

Jacqueline Wong
Erasmus University Rotterdam, the Netherlands
wong@essb.eur.nl

ABSTRACT: Massive Open Online Courses (MOOCs) suffer from low completion rates, questioning the potential benefits of open online education systems. To address the learning conundrum in MOOCs, the current project builds on three major theoretical concepts in the learning sciences, namely self-regulated learning (SRL), motivation, and cognitive load, to examine how the students regulate their learning in MOOCs and how they can be supported to enhance their use of SRL strategies. The ultimate goal is to enhance student engagement and success in MOOCs. The current paper gives an overview of the five studies and research methods that were proposed for this project. Learning analytic approaches will be used to analyze the large sets of data that will be collected in the experiments. The main challenge of the project is to model SRL processes by using different types of data, triangulating these data, and detecting actual SRL patterns from complex learning patterns.

Keywords: Self-regulated learning, motivation, student success, massive open online courses, learning analytics

1 INTRODUCTION

In the modern world, digital technology is an integral part of education for many countries (Selwyn, 2012). Report on the state of higher education online learning in the United States revealed a yearly increase in the number of students engaged in distance learning and more than 10% of the higher education institutes have Massive Open Online Courses, also known as MOOCs (Allen & Seaman, 2016). MOOCs can be broadly viewed as an evolving ecosystem of online learning environments (Rodriguez, 2012) offering free or low-cost education taught by an expert to anyone who is interested with no requirements on prior knowledge or attendance. The mainstream media discussed MOOCs as the future of education but soon the discussion turned into heavy criticism as studies showed that the attrition rate in an average MOOC is more than 90% (Jordan, 2014).

1.1 Challenges to learning in Massive Open Online Courses

The wide enrolment and completion gap raises concerns over the educational benefits of MOOCs. On one hand, research on student perspectives showed that reasons for dropping out include the lack of time, lack of incentives, lack of assistance, inability to comprehend materials, and other priorities (Hew & Cheung, 2014). This has been interpreted as not all students not have the necessary skills to regulate their own learning. The ability to regulate one's learning is crucial to achieving student success in MOOCs (Wong et al., 2017). On the other hand, research on instructional quality showed that MOOCs in general fare poorly on instructional design quality (Margaryan, Bianco, & Littlejohn, 2015). For instance, the MOOCs investigated did not support activation of prior knowledge nor application of new knowledge and skills. This suggests that research is needed to improve the instructional design quality of MOOCs. The lack of self-regulatory skills and high quality instructions combined drive the research agenda of the current project: How can students be supported to enhance study success in MOOCs?

2 THEORETICAL BACKGROUND

This project aims to integrate well-studied theories with documented impact on student success to develop a better understanding of how students regulate their learning and how they can be supported to become successful MOOC students. Figure 1 shows the conceptual model of the current project linking the theories that will be examined in the studies.

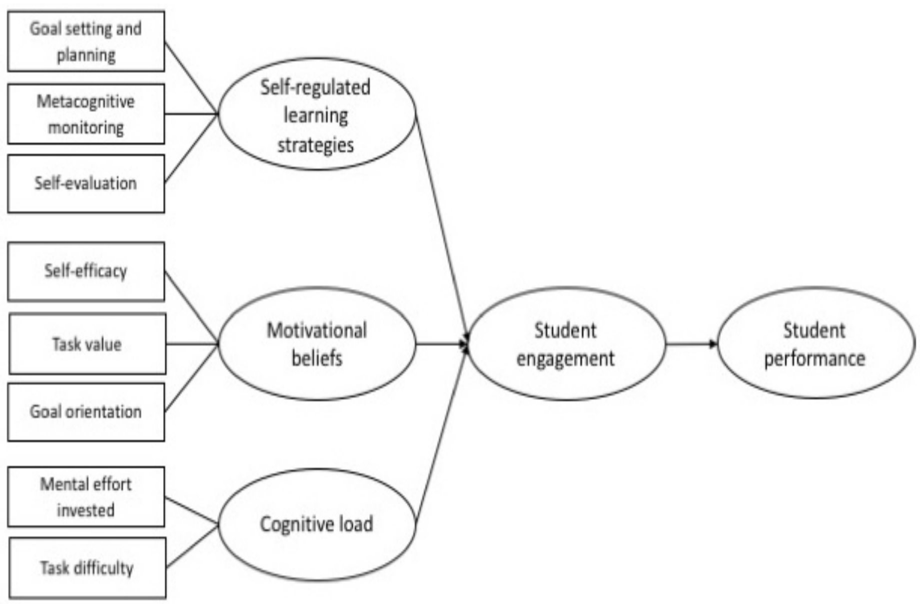


Figure 1: Conceptual model of the current project

2.1 Self-regulated learning

A systematic review by Broadbent and Poon (2015) showed that a positive correlation between grades in online higher education learning environments and self-regulated learning. Therefore, the results suggest that students who are not as well-versed in regulating their learning perform poorly or even give up learning early compared to students who are better at regulating their learning. Kizilcec, Pérez-Sanagustín, and Maldonado's (2017) study corroborated this finding as their results revealed that personal goal attainment in MOOCs was related to SRL, specifically goal setting and strategic planning. Moreover, students who reported higher levels of SRL spent more time revisiting assessments. Three main SRL strategies derived from the three phases of Zimmerman's (1989) SRL model are examined: a) goal-setting and planning in the forethought phase, b) metacognitive monitoring in the performance phase, and c) self-evaluation in the self-reflection phase.

2.2 Motivation

Motivational beliefs are essential to SRL (Winne, & Hadwin, 1998; Zimmerman & Campillo, 2003). Littlejohn, Hood, Milligan, and Mustain (2016) interviewed good and poor self-regulated students taking a MOOC and found differences in their description of motivations and goals for the course which in turn shaped their use of learning strategies. Therefore, to examine the effect of approaches to support SRL in MOOCs, it would be of interest to take into account students' motivation. Three motivational beliefs that are documented to influence learning are examined: a) self-efficacy, b) task value, and c) goal orientation.

2.3 Cognitive load

Research in Cognitive Load Theory (CLT) examines how instructional design can be designed to support learning of complex tasks by managing cognitive load (Paas, Renkl, & Sweller, 2003). To learn a complex task, students will have to self-regulate their learning by monitoring their performance during the task (Kostons, Van Gog, & Paas, 2012). However, the relationship between cognitive load and SRL has not been fully explored (de Bruin & van Merriënboer, 2017). It is not clear whether supporting SRL hinders learning by competing for the limited working memory resources or supports learning by reducing cognitive load.

3 RESEARCH METHODS

To develop an overview of the current approaches, two systematics reviews will be conducted. This project will also include a series of experiments to examine effects of various SRL supports.

3.1 Systematic Reviews

3.1.1 Review 1: State of the art approaches to support SRL

The goal of the systematic review was twofold: a) to inform researchers, designers, and teachers about the current approaches to support SRL in online learning environments and MOOCs, and b) to use the findings to design an SRL support for MOOCs. Therefore, two research questions were formulated:
RQ. R1a: What kind of approaches are there to support SRL?
RQ. R1b: Do student characteristics affect the efficacy of SRL supports?

Keywords related to SRL and online learning environments were used to search for relevant literature. Findings from the 35 studies reviewed showed that there was only one study (Kizilcec, Pérez-Sanagustín, & Maldonado, 2016) thus far that empirically examined an approach to support SRL in MOOCs. Overall, studies suggest that SRL support enhances the use of SRL strategies (e.g., Bannert & Reimann, 2012). Prompting has been identified as a promising approach to enhance SRL and learning performance (e.g., Sitzmann & Ely 2010). The current manuscript is in press. However, findings have been shared in recent conferences organized by the graduate school for educational sciences in the Netherlands and in the European Association for Research on Learning and Instruction.

3.1.2 Review 2: Bridging the gaps between learning sciences and learning analytics

Chatti, Dyckhoff, Schroeder, and Thüs (2012) proposed a four-dimensional reference model for learning analytics (LA) addressing the data and environment (what), the stakeholder (who), the objectives (why), and the techniques (how). Although it is a comprehensive model, it does not consider the role of learning theories in the LA process. For the analyses to be meaningful, the LA approach has to be framed within theoretical models of learning (Gašević, Dawson, & Siemens, 2015). Therefore, to close the gap between learning sciences and LA, a systematic review will be conducted to answer the research question:

RQ. R2: What are the LA approaches that have been used to examine constructs related to learning and how are the LA approaches employed?

3.2 Empirical studies to support SRL in MOOCs

3.2.1. Experiment 1: Prompting SRL using videos

Based on the findings in the first systematic review, the first experimental study examines the effect of using videos to prompt SRL in MOOCs. Students in the experimental group have access to videos that are embedded in the weekly modules of the course. Each video consists of three questions asking students to think of their current learning state in relation to planning, monitoring, and reflecting. The prompts are intended to activate the use SRL strategies when learning in MOOCs. The main research questions are:

RQ. E1a: What is the effect of using videos to prompt self-regulated learning strategies on student engagement and learning performances?

RQ. E1b: How do individual differences in SRL influence the effect of videos prompting SRL?

The study was conducted in three MOOCs offered by Erasmus University Rotterdam on the Coursera Platform, the MOOC provider. Three types of data were collected: 1) Self-report data using questionnaires, 2) clickstream data that consist of logs of students' interactions with the course materials, and 3) scores on quizzes and peer assignments. To model the process of SRL, the study will employ two learning analytics methodologies, sequential pattern analysis and social network analysis, to examine learning behavior of students in MOOC and how these behaviors relate to SRL.

3.2.1 Experiment 2: Exploring multimodal measurement of engagement and SRL

The second experiment will build on the first experiment by exploring the effect of training SRL to enhance the use of SRL strategies. With reference to the conceptual model of the current project described in the second section of this paper, multiple forms of measurements, such as electrodermal activity, eye tracking, log files, and self-reports will be used to examine the relationship between motivational beliefs, cognitive load, and SRL activities. Azevedo (2009) stressed the importance of using multiple forms of measurements to detect, track, and model how students use the different processes during SRL. By doing so, we can better understand how the processes individually or as a whole affect learning. In view of the explorative nature of the study and the logistics in deploying the multiple measurements, the study will be conducted in a lab using a repeated measures design. The two research questions are:

RQ. E2a: What is the effect of using videos to train self-regulated learning strategies on student engagement and learning performances?

RQ. E2b: How do the different processes measured by the multiple forms of measurements relate to each other and to student engagement and learning performances?

3.2.2 Experiment 3: Examining the effect of gamification on student motivation and SRL

The third experiment will involve designing of a mobile application that complements MOOCs to examine the effect of gamification to enhance motivation and SRL in MOOCs. Studies suggest that gamification enhances student engagement in online courses (Looyestyn et al., 2017). While gamification may take the form of progress bars in MOOCs, a major issue in MOOC research is the lack of compliance with the interventions. Preliminary findings from the first experiment showed that very few students interacted with the intervention. Therefore, instead of placing interventions within the MOOC environment where students would have a chance to access only when they log in, the third experiment will use a mobile application that can be accessed even when students are not logged on to the course. Therefore, it is likely that students will attend to the course more often with an intervention that can be easily accessed and uses gamification. Learning analytics approaches will be used in the experiment to create visualizations enabling students to monitor their learning as part of the game design. The research question for Experiment 3 is:

RQ E3: What is the effect of a gamified mobile application on student engagement and learning performances in MOOCs?

4 CONCLUSIONS

The scientific value of the project lies in the integration of theories of self-regulated learning, motivation, and cognitive load to advance our understanding of learning in the context of MOOCs. Moreover, the experimental studies conducted and proposed in this project will provide empirical evidence to evaluate the effectiveness of the above-mentioned approaches to support SRL. The studies in this project are novel as very few experimental studies examining SRL supports have been conducted in MOOCs. Adding to the level of originality is the exploration of various forms of process measures and learning analytics approaches to model students' SRL processes. These findings will set the stage for providing adaptive support based on different patterns of SRL. The interdisciplinary nature of the project creates opportunities for collaboration between researchers in computer science, pedagogy and psychology.

5 LIMITATIONS AND CHALLENGES

There are two major limitations in this project. First, selection bias is unavoidable in MOOCs because students have the choice to either interact or ignore the interventions. Second, the types of intervention that can be implemented are restricted by the options (e.g., text, open ended question or videos) on the MOOC platforms. Therefore, the project utilizes a mobile application in the third experiment to work around this restriction.

The biggest challenge of this project is to detect actual patterns of SRL from complex patterns of learning activities. This issue is exacerbated by the individual differences that influence SRL. Roll and Winne (2015) proposed collecting multiple forms of data as the solution. Increasing the types of data collected can possibly help us to identify meaningful learning patterns that indicate SRL.

REFERENCES

Allen, I. E., & Seaman, J. (2016). Online report card: Tracking online education in the United States. Retrieved from http://onlinelearningsurvey.com/reports/onlinereportcard.pdf

Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*, 193-211.

Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning, 4*, 87-95.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1-13.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5-6), 318-331.

de Bruin, A. B., & van Merriënboer, J. J. (2017). Bridging Cognitive Load and Self-Regulated Learning Research: A complementary approach to contemporary issues in educational research. *Learning and Instruction, 51*, 1-9.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*, 64-71.

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review, 12*, 45-58.

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distance Learning, 15*, 133-160.

Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education, 29*, 40-48.

Looyestyn, J., Kernot, J., Boshoff, K., Ryan, J., Edney, S., & Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PloS one, 12*(3), 1-19.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2016). Recommending self-regulated learning strategies does not improve performance in a MOOC. *In Proceedings of the Third (2016) ACM Conference on Learning@ Scale,* 101-104.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18-33.

Kostons, D., Van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22*, 121-132.

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77-83.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1-4.

Rodriguez, C. O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance and E-Learning*. Retrieved from http://www.eurodl.org/?p=Special&sp=init2&article=516

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics, 2*, 7-12.

Selwyn, N. (2012). Education in a digital world: Global perspectives on technology and education. Oxford: Routledge.

Sitzmann, T., & Ely, K. (2010). Sometimes you need a reminder: the effects of prompting self-regulation on regulatory processes, learning, and attrition. *Journal of Applied Psychology*, 95, 132-144.

Wong, J., Baars, M., Davis, D., Van der Zee, T., Houben, G. J., & Paas, F. (2017). Supporting Self-Regulated Learning in Online Learning Environments and MOOCs: A Systematic Review. Manuscript submitted for review.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*, 329-339.

Exploring the Impact of Personalized Feedback on Time Management Behaviour: A Learning Analytics Approach

Nora'ayu Ahmad Uzir
School of Informatics, University of Edinburgh
Faculty of Information Management, Universiti Teknologi MARA
n.uzir@ed.ac.uk

ABSTRACT: Recent research confirmed that ineffective time management and low self-regulation skills are the common combination of unsuccessful learning factors. This doctoral study aims to explore and provide empirical evidence how learning analytics-based feedback can promote effective time management behaviours and academic achievement. Specifically, the research looks at the ways how learning analytics can be used to i) generate personalized feedback based on students' learning activities and ii) analyze the effects of personalized analytics-based feedback on time management. This study will employ quantitative research methods consist of the Epistemic Network Analysis (ENA) and agglomerative hierarchical clustering. Besides, two self-report instruments will be used to complement the digital trace data which are Study Process Questionnaire R-SPQ-2F and Motivated Strategies for Learning Questionnaire (MSLQ). The results of this PhD research are expected to provide a socio-technical solution can contribute towards the provision of learning analytic based feedback at scale.

Keywords: Personalized feedback, learning analytics, self-regulated learning, time management

1 BACKGROUND

Time management is critically related to the notions of procrastination. Existing research indicates that a majority of learner population struggle to attain an appropriate level of self-regulation and suffers from frequent instances of procrastination (Yamada et al., 2016). Earlier research noted that procrastination is not only characterized by motivational issues like poor time management and traits of laziness, but it is also influenced by emotional failure and low self-regulation (Senécal, Koestner, & Robert, 1995; St, Macan, Dipboye, & Phillips, 1990). In essence, ineffective time management and low self-regulation skills are the most commonly cited combination of unsuccessful learning factors (Petersen, Craig, Campbell, & Tafliovich, 2016; Thibodeaux, Deutsch, Kitsantas, & Winsler, 2017)**.**

Over the past few decades, there has been a wealth of evidence that recognized feedback as a key factor to promote effective learning and overall academic achievement. Feedback is "information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way" (Ramaprasad, 1983, p. 4). As posited in the COPES model (Winne, 2013), learners evaluate their learning products and the effectiveness of learning strategies based on internal and external standards. The internal standards described as internal qualities of learners (e.g., experience, prior knowledge) that autonomously guide students' own learning (Winne, 2013). However, in some cases, students need to seek external feedback from an expert agent such as teachers, peers, or groups (Hattie & Timperley, 2007) to decrease discrepancies between their current learning state and intended learning outcomes (Butler & Winne, 1995).

Building from learning analytics and educational data mining research, learning dashboards has received much attention as a visualize feedback tool over the past few years. However, their effectiveness has been questioned (Pardo, 2017), learners often misunderstand and misinterpret the information presented in the dashboards (Corrin & de Barba, 2015; Ramos-Soto, Vázquez-Barreiros, Bugarín, Gewerc, & Barro, 2016) and failed to incorporate the feedback from dashboards to their learning strategies. According to Pardo, this is due to the lack of the instructor's knowledge in feedback provided by learning dashboards (Pardo, 2017). Another significant perspective is there is limited evidence on positive impact the usage of visualization on motivation, interpretation of data, and impact on learning success (Gašević, Kovanović, & Joksimović, 2017; Pardo, 2017).

Recent research recognized the importance of alignment of learning analytics with the literature on feedback to provide personalized feedback at scale. Considerable work has been demonstrated in the work by Pardo and colleagues on feedback provision (Pardo, 2017). According to Pardo and colleagues, the idea behind analytics-based personalized feedback is to combine digital trace data typically captured by computer mediated technology with the instructor knowledge to provide more elaborated and personalized feedback to individual learners in an instructional and timely manner (Pardo, Poquet, Martinez-Maldonado, & Dawson, 2017). The design of this approach includes the formation of feedback messages that are parametrized based on the indicators captured by learning analytics. This will then follow by an execution of an algorithm that will select suitable feedback options for individual learner based on their level of engagement with learning activities and send the message into their personal email. This process is repeated throughout an academic semester on a weekly basis.

Recent research efforts have confirmed the positive effects of analytics-based personalized feedback on student satisfaction and academic achievement (Pardo, Jovanovic, Dawson, Gasevic, & Mirriahi, in press). However, there has been a dearth of empirical studies on large quantitative investigations of learning analytics-based feedback that sought to promote effective time management behaviours in student learning. Explaining students' time management based on digital traces through learning analytics methods is expected to provide additional evidence towards better understanding of actual performance behaviours. Hence, this doctoral research aims to explore and provide empirical evidence how learning analytics-based feedback can promote effective time management behavior and academic achievement. To unlock the full potential of analytics-based feedback, these underexplored opportunities warrant for further investigation:

1.1 Feedback amount and duration

Feedback is a key to support productive self-regulated learning (Azevedo et al., 2013; Winne & Hadwin, 2013). Winne posits that learners require an optimal external support before they are able to gain their own cognitive footing (Winne, 1995). Therefore, a proper timing of a feedback intervention significantly influences the learning outcomes (Thornock, 2016). However, a majority of research on feedback timing is to study the quality of immediate and delayed feedback given to students to enhance learning process and performance (Shute, 2008). Thus, there seems to be a definite need for an exhaustive evaluation on the effects of feedback on student learning outcomes (Dawson, 2017; Pardo et al., 2017) to provide clear guidelines on the ideal timeframe to send feedback to learners (Attali & van der Kleij, 2017; Tanes, Arnold, King, & Remnet, 2011). In this study, I propose to investigate the desirable duration and amount of feedback should be given to the learners in order to improve their time management practice and academic performance.

1.2 Feedback message polarity

Both positive and negative polarity of feedback have an influential effect on learning experience and achievement, but this effect can be either encouraging or damaging (Hattie & Timperley, 2007). Positive feedback is powerful to increase a learner's motivation to learn by cueing the learner to continue using their existing strategies (Mitrovic, Ohlsson, & Barrow, 2013; Tanes et al., 2011). Inversely, the learners who constantly receive positive feedback may be less challenged, which can eventually lead to boredom (Garris, Ahlers, & Driskell, 2002). The positive side of negative feedback is to encourage students to correct specific behaviours, and thus suggest students to consider other strategies to achieve the desirable goal, however, constantly receiving negative feedback could make students feel incompetent and discouraged (Tanes et al., 2011). Therefore, feedback need to balance between positive and negative message to promote motivation rather than discouragement (Holmes & Papageorgiou, 2009; Weaver, 2006). Aligned with the concerns addressed by (Nixon, Brooman, Murphy, & Fearon, 2017; Pardo et al., 2017) this study aims to investigate on the overall tone of feedback message (e.g., positive, negative or both) that truly reflects learners' skills, especially time management skills.

2 AIM OF THE RESEARCH

The aim of this doctoral study is to gain deeper understanding on how learning analytics-based feedback can affects time management behaviour and academic achievement. To eliminate potential for bias and deficiencies introduced in standard self-report measures, this research focuses on the analysis of actual learner's behaviour based on log data in the Learning Management System (LMS) in the context of flipped learning courses. The trace data will be collected and analyzed by established data science methods and the results of the analysis will be interpreted based on self-regulated learning theory. In this study, Winne & Hadwin (1998) self-regulated learning theory is adopted as the theoretical framework due to its practical relevance for implementation (Gašević et al., 2017). Specifically, the purpose of this study is twofold: i) to analyze the effects of analytics-based personalized feedback on time management practices and ii) provide analytics-based personalized feedback for learners to improve their time management. The following questions will be used to guide the design of the research:
1: What are the effects of analytics-based feedback on learner's time management behaviours and academic performance?
2: Does and if so how the duration of analytics-based feedback affects time management? What is the optimal amount of analytics-based feedback should be given to learners to improve time management

and academic performance?
3: Does the type (negative or positive message) of analytics-based feedback can influence the time management behaviour and academic performance?

3 METHODOLOGY

This study will adopt quantitative research methods. Four data sources will be used to address the objectives of this study. First, trace data will be extracted from the Learning Management System (LMS) from four students' cohorts starting from year 2014 until 2017 in an engineering large enrollment (N 290 to 650 per cohort) undergrad course that followed a flipped classroom instructional design. In each year, each group of students have received different type of feedback interventions (e.g., learning dashboard or analytics-based personalized feedback) over different length of time (e.g., first-half of semester or throughout the semester) on a weekly basis. The second and third data source will be derived from midterm test scores and final examination scores. For data analysis, Epistemic Network Analysis (ENA) (Shaffer, Collier, & Ruis, 2017) will be used to analyze data from LMS together with midterm and final examinations scores to address the effects of feedback on time management practice and academic performance for four different cohorts. Agglomerative hierarchical clustering will be adopted to identify different groups of students based on types of feedback (positive of negative feedback) they received. Next, performance of each groups in midterm and final examination will be compared to examine the effectiveness of different type of feedback. Finally, two self-report instruments will be used to complement the digital trace data which are Study Process Questionnaire R-SPQ-2F (Biggs, Kember, & Leung, 2001) and Motivated Strategies for Learning Questionnaire (MSLQ) (Duncan & McKeachie, 2005). These two self-reports were administered at the beginning of each course in order to identify the time management aspects of the students learning.

4 CURRENT STATUS AND RESULTS ACHIEVED

Progress to date is twofold: The first study is a systematic review of the literature of learning analytics dashboard research through the lens of self-regulated learning theory by Winne & Hadwin (1998). This study aimed to explore and to provide evidence based on empirical studies of learning dashboards in supporting self-regulated learning. During the initial search work, a total of 382 papers were obtained from five main academic databases and Google Scholar. After a systematic assessment was conducted, a total of 29 empirical articles were included in the final analysis. During the paper analysis, each indicator presented in learning analytics dashboards were coded according to five elements in COPES model namely condition, operation, product, evaluation and standard. The result (Figure 1) indicated that most of the indicators reported on personal qualities of learners (e.g., prior knowledge), motivational aspects (e.g., highest score), and operations performed by learners during the learning process (e.g., number of login). It is noteworthy that an unequal attention was paid to different elements of the COPES model, particularly there is little attention paid to task conditions in the existing literature. This study concluded that the available learning dashboards were not well-grounded in an established educational theory (Gašević, Dawson, & Siemens, 2015). Therefore, feedback presented through currently available learning dashboards is unlikely to yield useful and actionable recommendation to help individual learner on their metacognitive strategies or time management strategies. This gap offers starting point for further research on learning analytics-based feedback to promote self-regulation in students learning and better time management.

The second study investigated learner interaction with online preparatory tasks in a flipped classroom in terms of time management behaviour and learning strategies. The traced data collected from 290 first year students enrolled in the Computer Systems course collected during the course offering in 2014. The finding revealed some interesting self-regulations patterns in term of time management through comparing between high and low performance students. Note that, low performance score was cut off at 25 percentiles whereas high performance was cut off at 90 percentiles for both midterm test and final examination as per Jovanovic and her colleagues (2017). The results showed that (refer to Figure 2 and Figure 3): i) low achievers tended to exhibit more procrastination behaviour by not completing the required tasks on time and by delayed access to learning materials and ii) low achievers were characterized by inconsistency in time management in which they tried to be ahead of time in visiting particular topics but failed to maintain it in the rest of the course. It can be concluded that learners spent their time in a way that did not support their learning, which could have an undesirable impact on their learning outcomes.

Next step, these two papers will be sent out for publication. In term of ongoing research, log data from year 2015 until 2017 will be analyzed to provide insights into the effects of different feedback interventions on time management practice and academic performance.

REFERENCES

Attali, Y., & van der Kleij, F. (2017). Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. *Computers and Education*, *110*, 154–169. https://doi.org/10.1016/j.compedu.2017.03.012

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-behnagh, R., Bouchet, F., & Landis, R. (2013). International Handbook of Metacognition and Learning Technologies, *28*, 427–449. https://doi.org/10.1007/978-1-4419-5546-3

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The Revised Two Factor Study Process Questionnaire : R-SPQ-2F The Revised Two Factor Study Process Questionnaire : R-SPQ-2F. *British Journal of Educational Psychology*, *71*, 133–149. https://doi.org/10.1348/000709901158433

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, *65*(3), 245–281. https://doi.org/10.2307/1170684

Corrin, L., & de Barba, P. (2015). How Do Students Interpret Feedback Delivered via Dashboards? *International Conference on Learning Analytics and Knowledge*, 430–431. https://doi.org/10.1145/2723576.2723662

Dawson, S. (2017). From prediction to impact : Evaluation of a learning analytics retention program, 3–7. https://doi.org/10.1145/3027385.3027405

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, *40*(2), 117–128.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming*, *33*(4), 441–467. https://doi.org/10.1177/1046878102238607

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1). https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the Learning Analytics Puzzle: A Consolidated Model of a Field of Research and Practice, (January). https://doi.org/10.1080/23735082.2017.1286142

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Holmes, K., & Papageorgiou, G. (2009). Good, bad and insufficient: Students' expectations, perceptions and uses of feedback. *The Journal of Hospitality Leisure Sport and Tourism*, *8*(1), 85–96. https://doi.org/10.3794/johlste.81.183

Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*. https://doi.org/10.1016/j.iheduc.2017.02.001

Mitrovic, A., Ohlsson, S., & Barrow, D. K. (2013). The effect of positive feedback in a constraint-based intelligent tutoring system. *Computers and Education*, *60*(1), 264–272. https://doi.org/10.1016/j.compedu.2012.07.002

Nixon, S., Brooman, S., Murphy, B., & Fearon, D. (2017). Clarity, consistency and communication: using enhanced dialogue to create a course-based feedback strategy. *Assessment and Evaluation in Higher Education*, *42*(5), 812–822. https://doi.org/10.1080/02602938.2016.1195333

Pardo, A. (2017). A feedback model for data-rich learning experiences. *Assessment and Evaluation in Higher Education*, *2938*(October), 1–11. https://doi.org/10.1080/02602938.2017.1356905

Pardo, A., Poquet, O., Martinez-Maldonado, R., & Dawson, S. (2017). Provision of Data-Driven Student Feedback in LA & EDM. *Handbook of Learning Analytics*, 163–174. https://doi.org/10.18608/hla17.014

Petersen, A., Craig, M., Campbell, J., & Tafliovich, A. (2016). Revisiting why students drop CS1. *Proceedings of the 16th Koli Calling International Conference on Computing Education Research - Koli Calling '16*, 71–80. https://doi.org/10.1145/2999541.2999552

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*(1), 4–13. https://doi.org/10.1002/bs.3830280103

Ramos-Soto, A., Vázquez-Barreiros, B., Bugarín, A., Gewerc, A., & Barro, S. (2016). Evaluation of a Data-To-Text System for Verbalizing a Learning Analytics Dashboard. *International Journal of Intelligent Systems*, (May). https://doi.org/10.1002/int.21835

Senécal, C., Koestner, R., & Robert, J. (1995). Self-regulation and academic procrastination. *The Journal of Social Psychology*, *135*(5), 607–619. https://doi.org/http://dx.doi.org/10.1080/00224545.1995.9712234

Shute, V. J. (2008). Focus on Formative Feedback. *Source: Review of Educational Research*, *78228173*(1), 153–189. https://doi.org/10.3102/0034654307313795

St, M., Macan, T. H., Dipboye, R. L., & Phillips, A. P. (1990). College Students â€™ Time Management : Correlations With Academic Performance and Stress. *Journal of Educational Psychology*, *82*(4), 760–768. https://doi.org/10.1037/0022-0663.82.4.760

Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers and Education*, *57*(4), 2414–2422. https://doi.org/10.1016/j.compedu.2011.05.016

Thibodeaux, J., Deutsch, A., Kitsantas, A., & Winsler, A. (2017). First-Year College Students' Time Use. *Journal of Advanced Academics*, *28*(1), 5–27. https://doi.org/10.1177/1932202X16676860

Thornock, T. A. (2016). How the timing of performance feedback impacts individual performance. *Accounting, Organizations and Society*, *55*, 1–11. https://doi.org/10.1016/j.aos.2016.09.002

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, *31*(3), 379–394. https://doi.org/10.1080/02602930500353061

Winne, P. H. (2013). Learning strategies, study skills, and self-regulated learning in postsecondary education. In *Higher education: Handbook of theory and research* (pp. 377–403). Springer.

Winne, P. H., & Hadwin, A. F. (1998). Studying as Self-Regulated Learning. *Metacognition in Educational Theory and Practice*.

Winne, P. H., & Hadwin, A. F. (2013). International Handbook of Metacognition and Learning Technologies, *28*, 293–308. https://doi.org/10.1007/978-1-4419-5546-3

Yamada, M., Goda, Y., Matsuda, T., Saito, Y., Kato, H., & Miyagawa, H. (2016). How does self-regulated learning relate to active procrastination and other learning behaviors? *Journal of Computing in Higher Education*, *28*(3), 326–343. https://doi.org/10.1007/s12528-016-9118-9

APPENDIX

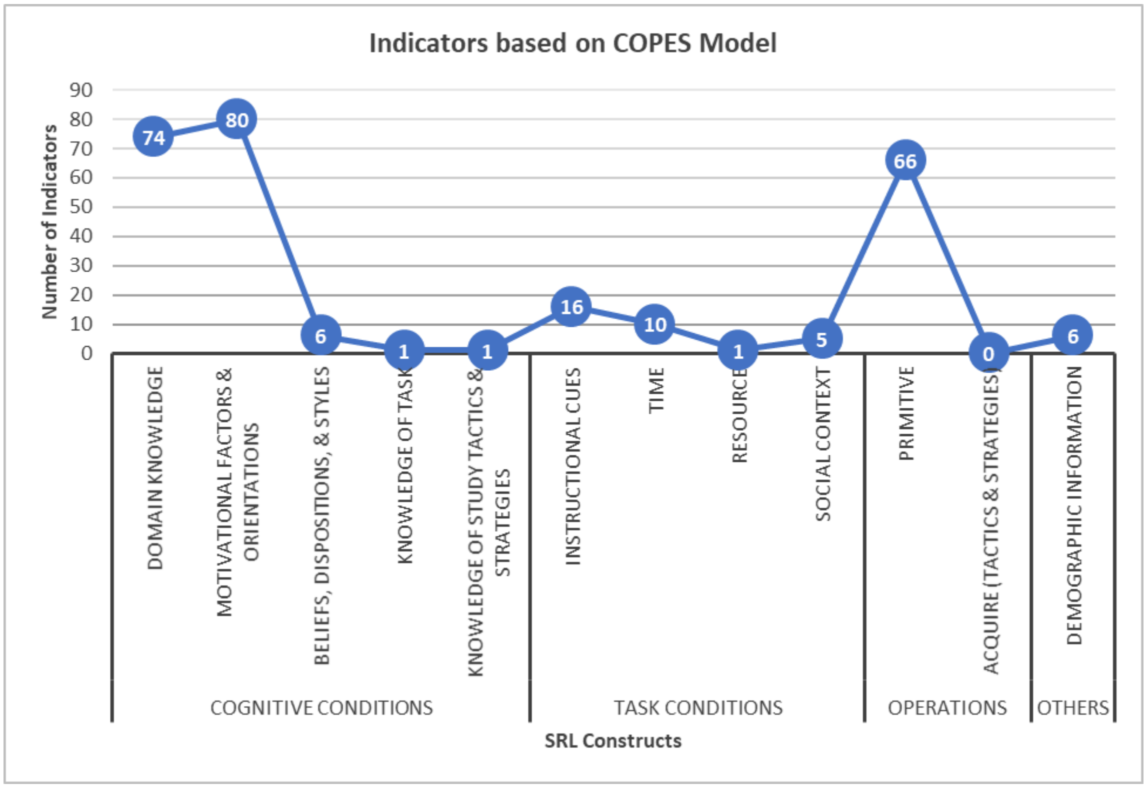


Figure 1: Number of indicators according to COPES elements and learning phases that the indicators were captured.

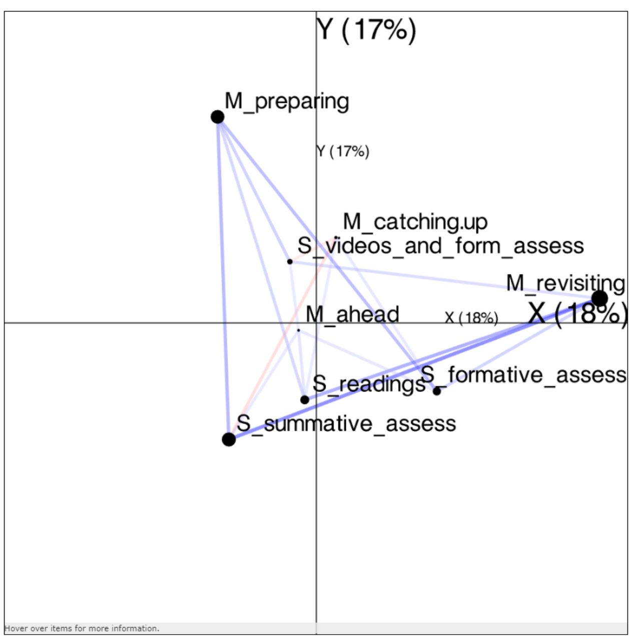


Figure 2: Comparison of high performance (blue line) and low performance (red line) students in midterm test according to their time management and learning strategies by using Epistemic Network Diagram (ENA)

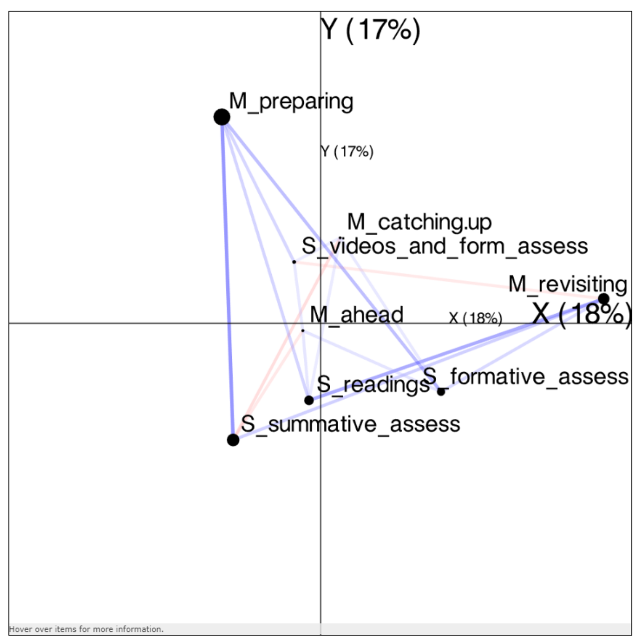


Figure 3: Comparison of high performance (blue line) and low performance (red line) students in final examination according to their time management and learning strategies by using Epistemic Network Diagram (ENA)

Beyond Data Presentation: Learning Analytics to Uncover Learning Strategies and the Influences of Feedback

Wannisa Matcha
School of Informatics, University of Edinburgh
w.matcha@ed.ac.uk

ABSTRACT: Ability to effectively regulate own learning has become one of the most important skills in the digital age. Effective self-regulated learners apply potent learning strategies. Feedback is a driving force behind the selection of learning strategies, but self-generated feedback is often not accurate. Therefore, suitable feedback provided by external agents can guide the adoption of more effective learning strategies. However, there is no holistic understanding on how to administer feedback that can promote potent learning strategies at scale. Moreover, current methods used to capture learning strategies based on self-reports are susceptible to different forms of bias. By applying learning analytics method to analyse trace data collected from the Learning Management System in real class settings, this doctoral research aims to produce evidence on how students unfold their learning strategies over time and how external analytics-based feedback influences their learning. This doctoral research also aims to provide a guideline on how to address the feedback to promote positive influences on learning strategy and therefore enhance the self-regulated learning process.

Keywords: Learning analytics; feedback; self-regulated learning;

1 INTRODUCTION

Self-regulation of own learning is probably one of the most crucial learning skills. However, students are not always successful in regulating their learning (Kizilcec, Pérez-Sanagustín, & Maldonado, 2016). Learning strategy is a key in self-regulated learning. High-performance learners employ different learning strategies than those who are considered to struggle in learning and who have low performance (DiFrancesca, Nietfeld, & Cao, 2016; Proctor, Prevatt, Adams, Reaser, & Petscher, 2006). Understanding and reinforcing effective learning strategy could enhance self-regulated learning skills. Feedback is an important factor in learning. According to Butler and Winne (1995), feedback is “an inherent and primer determiner of self-regulated learning” (p. 245). Providing students with suitable feedback can positively affect learning performance and selecting of effective learning strategies. However, poor feedback could negatively impact learning such as demotivating students. Despite the importance of feedback, the time constrains and workload create a challenge for teaching staff to provide timely and accurate feedback to their students. Nonetheless, there is no holistic guideline on how to administer feedback (Pardo, Jovanovic, Dawson, Gasevic, & Mirriahi, 2017). This research aims to fulfil this gap.

2 BACKGROUND

Winne and Hadwin (1998) posit that self-regulated learning process involves four cyclical recursive phases, including, task definition, goal and planning, enhancement of strategies and tactics, and adaptation. They propose a model that illustrates the key elements in each learning phase namely, Conditions, Operations, Products, Evaluate, and Standards (COPES). In early phases of learning, students define tasks and set learning goals by considering internal and external conditions. Examples of internal conditions can be knowledge on tasks, domain topic, learning tactics, and different facets of motivation. External conditions can be resources available, learning environment, instructional cues, and time constraints to complete a given task. Based on these conditions, students make judgement to set learning goals, plan their schedule, and set some expectation or 'standards'. Students 'operate' their learning by applying the selected learning strategies and tactics. At the end, the 'products' of learning are produced. Students 'evaluate' these products and the choices of learning strategies against the standards that had earlier set. This evaluation can result in the change of learning strategies or the updating of the standard (Winne & Hadwin, 1998)

Learning strategy is denoted as the factors that predict success in learning (Winne, 2006). Weinstein, Husman, & Dierking, (2000) define learning strategy as "any thoughts, behaviors, beliefs or emotions that facilitate the acquisition, understanding and later transfer of new knowledge and skills" (p.227). Research findings reveal that not all students applied the effective learning methods (Dunlosky, 2013; Malmberg, Sanna, & Kirschner, 2014). Moreover, there are differences in learning strategy application between low and high-performance students (DiFrancesca et al., 2016; Proctor et al., 2006). Therefore, there is a need to guide students towards the selection of effective learning strategies (Dunlosky, 2013; Winne, 2013). Research recognizes learning strategy from different dimensions from resource management (Bos & Brand-Gruwel, 2016), time management (Kizilcec et al., 2016), learning tactics (DiFrancesca et al., 2016), learning processes (Jovanovic, Gasevic, Dawson, Pardo, & Mirriahi, 2017), and tools selection (Gasevic, Mirriahi, Dawson, & Joksimovic, 2017).

Feedback is one of the most influential drivers in learning, especially in self-regulation process (Butler & Winne, 1995). Feedback influences student's beliefs, self-efficacy, motivation, standards, conditions and consequently impact on choices of learning strategy students apply in learning. Students make judgement on how to proceed with the tasks based on internal and external conditions and feedback received. The perceive of learning progress and performance feedback affects the selection of learning strategy (Malmberg et al., 2014; Winne & Perry, 2000). However, Bjork, Dunlosky, & Kornell, (2013) state that students are inaccurate in judging their own performance which can lead to the selection of an ineffective learning strategy (Malmberg et al., 2014).

Despite the importance of feedback, the time constrains and workload create a challenge for teaching staff to provide timely and accurate feedback to student (Pardo et al., 2017; Price, Handley, Millar, & O'Donovan, 2010). Taking advantage of technology and analytical methods, learning analytics can offer insights into learning by analysing trace data. Learning analytics could also support the provision of feedback. Generally, there are two approaches that are being used to address the feedback in learning analytics field namely, learning analytics dashboards and personalised feedback. Learning analytics dashboards build on the assumption that visualization help to increase the insights about learning process (Khan & Pardo, 2016; Schwendimann et al., 2016). Whereas, personalised feedback takes into account individual differences and provides specific feedback, usually in a textual format, to a specific student based on their performance and factors identified.

Feedback, and especially when presented in the form of learning analytics dashboards, requires interpretation (Price et al., 2010). As highlighted by much existing research, students who receive feedback sometimes do not understand it, not able to make use of it or recognise benefits from it (Evans, 2013). Some student perceives the feedback as not applicable or irrelevant to their ongoing tasks (Price et al., 2010). In addition, there is a limited understanding in degree of feedback that can be provided in dashboard. Nonetheless, much is still not understood on how to effectively provide this type of feedback. Imprudent provision of dashboard-based feedback can introduce the "detrimental instructional practice" (Gašević, Dawson, & Siemens, 2015, page. 69). In addition, there is no holistic understanding on how to provide effective feedback e.g. the property of feedback, the frequency that influent learning process and outcome (Hattie & Timperley, 2007). Moreover, using analytics technology to provide real-time feedback is under-explored area (Gašević et al., 2015; Pardo et al., 2017; Schwendimann et al., 2016). Also, the existing guideline does not take into account rich information produced by learning analytics which extracted from the actual behavior and patterns of students learning extracted from digital trace data (Pardo, 2017). Also, research on feedback in general often neglects individual differences (Malmberg et al., 2014).

Capturing how feedback influence the change in learning strategy is a difficult task because learning strategy is "latent construct" which is sometimes invisible and difficult to be observed (Jovanovic et al., 2017). Much research into student learning strategies has relied on self-reports that are collected through a questionnaire or think-aloud protocols (Pardo et al., 2017; Winne, 2013, 2014). Students' memories about the choices of learning strategies are often biased and incomplete (Winne, 2013), whereas using think-aloud protocol could impede students' learning due to cognitive overload (Winne, 2014). Data collected from digital environment such as log data reveals the actual behavior and learning pattern are not fully utilized (Siemens, 2013). This research aims to fulfill this gap by using data mining techniques and learning analytics methods to analyse log data, provide insights into how students learn.

3 RESEARCH QUESTIONS

The overall goal of this research is to contribute new insights into how learning analytics-based feedback can be provided effectively to guide students and to promote the adaptation of good learning strategies. the following research questions are formulated to guide the direction of this research:

RQ1: What and how have learning analytics approaches been used to support self-regulated learning?

RQ2: What are the effects of learning analytics-based feedback on the adoption of learning strategies and learning achievement?

RQ3: To what extent are learning strategies adopted aligned with learning analytics-based feedback provided?

RQ4: Does and if so how the time of learning analytics-based feedback influence the use of learning strategies?

RQ5: What is the effect of incorporation of motivational messages into learning analytics-based feedback on learning strategy adoption and performance?

Based on these research questions, we aim to promote the application of personalised feedback that takes advantage of learning analytics research by considering the actual behaviour contributed by a large cohort of students during their learning.

4 METHODOLOGY

The methods used in this study are in twofold. Firstly, research into the role of learning analytics begins with a literature review on learning analytics in order to find out the state of art in the area. Schwendimann et al. (2016) provide a systematic review of the literature on learning analytics dashboards. However, their work did not analyse the papers based on any grounded educational theory, which is essential if we aim to understand the extent to which existing learning analytics dashboards affect learning processes. Therefore, a systematic literature review has been conducted grounded around the SRL theory proposed by Winne and Hadwin (1998).

Secondly, to explore the effects of learning analytics-based feedback on the adoption of learning strategies and learning achievement, the analysis of trace data is explored. Four years of trace data will be used in this research. The data were collected from Learning Management System (LMS) based on flipped classroom (N=290-650). In each year, students received different type of feedback interventions. In 2014, feedback was addressed by using dashboards. In 2015, student received learning analytics-based personalized feedback during the first half of the semester and they could also view dashboard throughout the semester. In 2016, students were provided with personalized feedback and dashboard throughout the semester. In 2017, personalized feedback which was incorporated with motivational message and dashboard were provided throughout the semester. In addition, two self-report instruments namely, Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich & Groot, 1990) and Study Process Questionnaire (SPQ) (Biggs, 1978) will be used to complement the digital trace data. The self-reports were collected in the beginning of each course offering.

To explore the influences of feedback, the analysis begins with coding the sequences of learning activities into the corresponding learning activities and learning strategies. Epistemic Network Analysis (ENA) will be applied to explore how students unfold their learning strategies in each week. The comparison of ENA networks on different years of dataset will be carried out to probe the differences in learning strategies adopted by students who received different feedback interventions across the years. Additionally, the comparison of ENA networks of low and high-performance students will be conducted in order to discover the influences of learning analytics based feedback on learning performance.

Exploring how different type of feedback message (i.e. positive and negative feedback) influence (if at all) the change in learning strategies over time can be beneficial for informing practice for the use of learning analytics-based feedback. The agglomerative hierarchical clustering will be used to group the students based on the type of feedback message received (positive and negative feedback). ENA can be used to explore the differences of learning strategies adopted by students who received different type of feedback. Moreover, process mining will also be used to discover the common learning process of each group of students. The inferential statistics will then be used to examine the association of motivational factors retrieved from questionnaires and learning performance.

5 CURRENT STATUS & RESULTS ACHIEVED

To date, two studies have been carried out. Firstly, the systematic literature review, addressing RQ1, showed that research on learning analytics dashboards often neglect educational theory to inform the design and development of dashboards. Most of the dashboards used visualizations as a form of feedback about learning progress. Students are required to interpret such visualizations and make their own interpretations of the visualizations. There have been a small number of dashboards that offer recommendations about further actions needed to be taken by students in order to overcome or improve learning performance. There was no dashboard that recommended learning strategy to students to take. Outputs affected by learning analytics dashboards were varied. Some studies reported the perceived usefulness of dashboard where some reported otherwise. For example, Khan & Pardo (2016) found no significant associated between a number of dashboard view and midterm score.

Secondly, the trace data analysis was conducted by using ENA to explore how student unfold their learning strategies and how low and high-performance students apply their learning strategies (N=290), as a methodological preparation to address research questions 2-5. The learning patterns were observed (see Appendix). Figure 1 (in Appendix) reveals the common pattern observed during week 3-5. In general, students emphasized the use of summative assessment opportunities in preparatory activities and later revisiting those opportunities. The revisiting through self-testing is considered one of the effective learning strategy (DiFrancesca et al., 2016; Dunlosky, 2013). The differences between high and low performing students was also observed in the frequency of learning activities completed (Hattie & Timperley, 2007). Those who performed better on their mid-term and final exams showed a high frequency interacting with almost every learning section as presented in Figure 2 (see Appendix). As stated by Hattie & Timperley, (2007) increasing efforts reflect the productive learning activities. Moreover, a recent study in online learning environments also reports that individuals with strong self-regulated learning skill exhibits frequent revisiting action after completing learning material rather than immediately start with new material (Kizilcec et al., 2016).

REFERENCES

Bjork, R. a., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*(1), 120928131529005. http://doi.org/10.1146/annurev-psych-113011-143823

Bos, N., & Brand-Gruwel, S. (2016). Student differences in regulation strategies and their use of learning resources. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 344–353. http://doi.org/10.1145/2883851.2883890

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, *65*(3), 245–281.

DiFrancesca, D., Nietfeld, J. L., & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables. *Learning and Individual Differences*, *45*, 228–236. http://doi.org/10.1016/j.lindif.2015.11.010

Dunlosky, J. (2013). Strengthening the Student Toolbox. *American Educator*, *37*(3), 12–21. Retrieved from http://www.aft.org/sites/default/files/periodicals/dunlosky.pdf

Evans, C. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research*, *83*(1), 70–120. http://doi.org/10.3102/0034654312474350

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1). http://doi.org/10.1007/s11528-014-0822-x

Gasevic, D., Mirriahi, N., Dawson, S., & Joksimovic, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior*, *67*, 207–220. http://doi.org/10.1016/j.chb.2016.10.026

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. http://doi.org/10.3102/003465430298487

Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*, 74–85. http://doi.org/10.1016/j.iheduc.2017.02.001

Khan, I., & Pardo, A. (2016). Data2U : Scalable Real time Student Feedback in Active Learning Environments. *LAK '16 6th International Conference on Learning Analytics and Knowledge, April 25 - 29, 2016*, 249–253. http://doi.org/10.1145/2883851.2883911

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2016). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*. http://doi.org/http://dx.doi.org/10.1016/j.compedu.2016.10.001

Malmberg, J., Sanna, J., & Kirschner, P. A. (2014). Elementary school students' strategic learning: Does task-type matter? *Metacognition and Learning*, *9*(2), 113–136. http://doi.org/10.1007/s11409-013-9108-5

Pardo, A. (2017). A feedback model for data-rich learning experiences. *Assessment and Evaluation in Higher Education*, *2938*(October), 1–11. http://doi.org/10.1080/02602938.2017.1356905

Pardo, A., Jovanovic, J., Dawson, S., Gasevic, D., & Mirriahi, N. (2017). Using Learning Analytics to Scale the Provision of Personalised Feedback. *British Journal of Educational Technology*.

Pintrich, P. R., & Groot, E. V. De. (1990). Motivational and Self-Regulated Learning Components of Classroom Academic Performance. *Journal of Educational Psychology*, *82*(1), 33–40.

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback : all that effort , but what is the effect ? *Assessment & Evaluation in Higher Education*, *35*(3), 277–289. http://doi.org/10.1080/02602930903541007

Proctor, B. E., Prevatt, F. F., Adams, K. S., Reaser, A., & Petscher, Y. (2006). Study Skills Profiles of Normal-Achieving and Academically-Struggling College Students. *Journal of College Student Development*, *47*(1), 37–51.

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., … Dillenbourg, P. (2016). Perceiving learning at a glance : A systematic literature review of learning dashboard research, *X*(c), 1–14. http://doi.org/10.1109/TLT.2016.2599522

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. http://doi.org/10.1177/0002764213498851

Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning

theory perspective. *Educational Research Review*, *9*, 1–15. http://doi.org/10.1016/j.edurev.2012.11.004
Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. *Handbook of Self-Regulation*, *22*, 727–747.
Winne, P. H. (2006). How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist*, *41*(1), 5–17. http://doi.org/10.1207/s15326985ep4101_3
Winne, P. H. (2013). Learning Strategies, Study Skills, and Self-Regulated Learning in Postsecondary Education. *Higher Education: Handbook of Theory and Research*, (28), 337–403. http://doi.org/10.1007/978-90-481-8598-6
Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, *9*(2), 229–237. http://doi.org/10.1007/s11409-014-9113-3
Winne, P. H., & Hadwin, A. F. (1998). Studying as Self-Regulated Learning. *Metacognition in Educational Theory and Practice*.
Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In *Handbook of Self-Regulation* (pp. 531–599).

APPENDIX

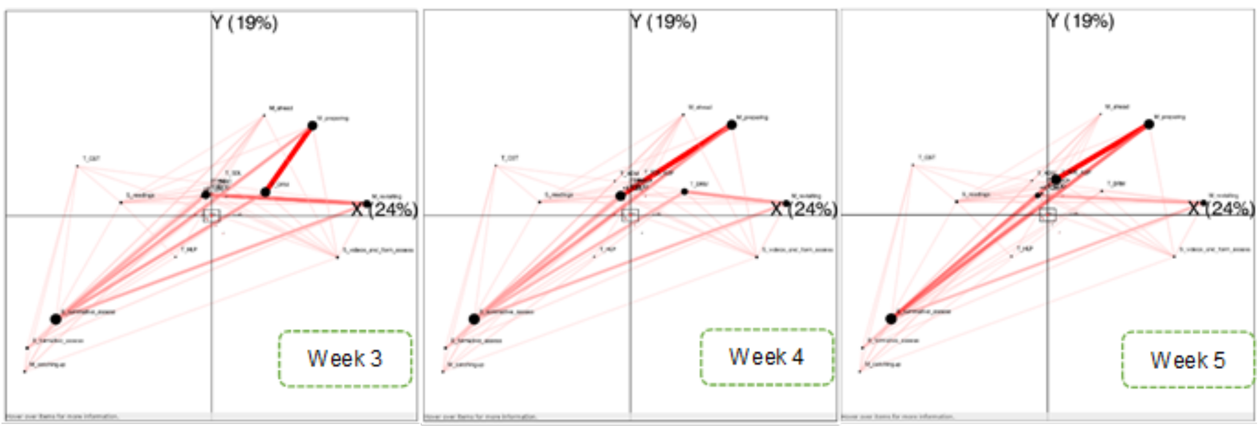


Figure 1: Learning patterns observed with ENA (during week 3, 4 and 5) illustrates the emphasise on the use of summative assessment (problem-solving activities and revisiting)

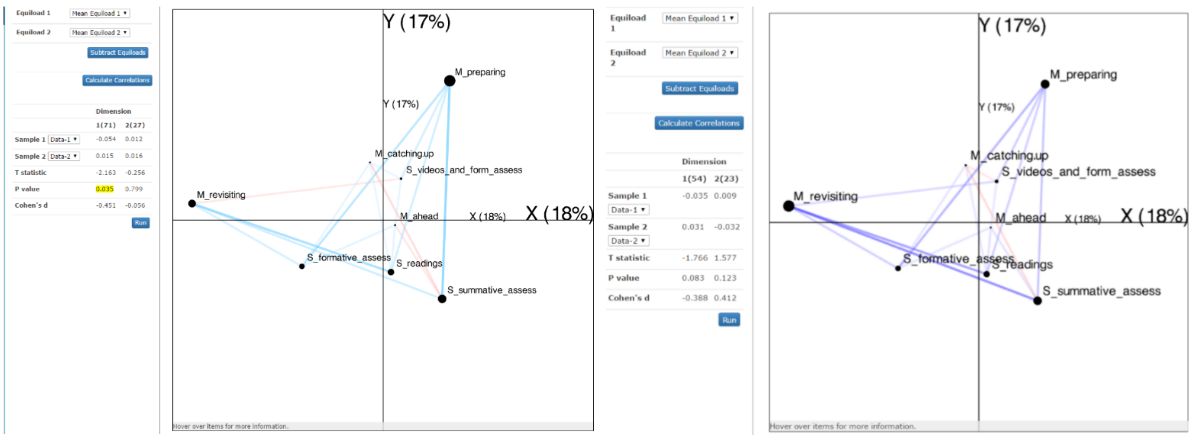


Figure 2: Learning pattern generated from ENA (Left: mid-term score and Right: final exam score) illustrates the differences between low-performance (red line) and high-performance (blue line)

The use of complex adaptive systems and meso-level practitioners to enhance and transform learning analytics implementation capability

Author: Colin Beer
University of Southern Queensland
colinwbeer@gmail.com

ABSTRACT: Universities are struggling to develop the capabilities required to implement meaningful learning analytics. Approaches to learning analytics that overlook the individual nature of learning and the complexity of universities as organisations are contributing to this problem. Human meaning-making and attention to variables in specific contexts is required for meaningful learning analytics. These notions contrast with university strategies that are dominated by best practice, vanilla implementations and centralised approaches to technology adoption. This study aims to develop design principles that will help universities cultivate the learning analytics implementation capabilities required to bridge the gap between centralised approaches and the individual needs of teachers and students.

Keywords: complexity, complex adaptive systems, learning analytics, implementation

1 BACKGROUND TO THE PROJECT AND THE IDENTIFICATION OF SIGNIFICANT PROBLEMS

The rapid increase in participation rates across the Australian higher education sector is contributing to the proliferation of online education. This has created a challenge for universities, as online students are not directly observable by the university in terms of their engagement with their studies when compared with the traditional classroom. Universities rely on technologies to provide insight into student progress and engagement. While the current generation of university information systems collect vast amounts of detailed data on staff and student behaviour, the tools that are provided to utilise this data are rudimentary at best (Motz, Teague, & Shepard, 2015). While learning analytics is touted as a potential solution to this problem, universities are struggling to develop the required capabilities to embed learning analytics across disparate learning and teaching contexts (Colvin, Dawson, Wade, & Gasevic, 2017).

Learning is a complex social activity (Siemens & Baker, 2012) situated in a complex social environment (Macfadyen & Dawson, 2012). The addition of vast amounts of detailed data into these complex and diverse environments makes it challenging for organisations to develop actionable intelligence suitable for different contexts (Clow, 2014). This difficulty is exacerbated by the strategic approach that universities apply to emerging technology related concepts like learning analytics (D. T. Jones & Clark, 2014). Single, integrated and centralised technologies are deemed to be efficient and are implemented according to detailed plans designed to achieve a pre-identified future state (D. T. Jones & Clark, 2014). Yet learning analytics is an multifaceted construct with many interdependent and contributing factors (Colvin et al., 2017).

The misalignment between the complex nature of learning analytics implementations and the strategic operations of universities has the potential to challenge and disrupt traditional management and organisational structures (Colvin et al., 2017). The incompatibility between exclusively top-down or bottom-up approaches to learning analytics implementation remains an unresolved challenge for universities. Top-down approaches will be unlikely to meet the equivocal needs of specific learning and teaching contexts whereas bottom-up approaches conflict with organisational expectations around centralisation and efficiency. Considering learning analytics through an alternative conceptual lens may help universities address this issue and move learning analytics beyond its current status as an unrealised aspiration (Ferguson et al., 2014).

Universities are complex socio-technical systems (Macfadyen, Dawson, Pardo, & Gasevic, 2014). Other socio-technical systems, most notably healthcare, have applied an alternative ontological conceptualisation based on complex adaptive system theory in an effort to move beyond hierarchical and mechanical models (Boustani et al., 2010; Plsek & Greenhalgh, 2001) . This theory describes systems that are non-causal, non-linear and are comprised of many interacting and interdependent agents (Holland, 2006). Complexity leadership and distributed leadership models are related to complex adaptive systems and can potentially contribute to learning analytics implementation and uptake (Colvin et al., 2017). The application of a complex adaptive systems conceptualisation in universities shifts the epistemological approach from planning and strategy, to an approach based on learning and improvisation (Juarrero, 1999; Kurtz & Snowden, 2003). While an alternative conceptualisation of learning analytics may assist with implementation and uptake, it must also fit within the current strategic and hierarchical operating norms of universities.

Considering learning analytics as either top-down or bottom-up runs the risk of overlooking a key area of translation in between the two (Hannon, 2013). Much of the work of implementation occurs between the top-down and bottom-up level, at the meso-level. This is the level within an organisation that sits between the small scale, local interactions and the large-scale policy and institutional processes (C. Jones, Dirckinck - Holmfeld, & Lindström, 2006). The practitioners who operate at the meso-level are crucial agents who translate between the centralised technologies and strategies at the upper levels, and the contextual specificities at the lower levels. These meso-level practitioners assuage the tension between the upper and lower levels (Uhl-Bien, Marion, & McKelvey, 2007). The contextual nature of learning analytics and the centralised approach to data services common to universities make the meso-level practitioners a crucial element in learning analytics implementation.

2 GOALS OF THE RESEARCH AND A CLEAR FORMULATION OF THE RESEARCH QUESTION

This project aims to produce design principles derived from complex adaptive systems theory to guide the contribution of meso-level practitioners in order to help universities address the challenges of institutional learning analytics implementation. The principles aim to improve the integration between tools, actionable data and educator practices, in real-word settings. The principles will be iteratively developed and refined through a cycle of design-based research within a regional Australian university with the broad aim of improving student learning outcomes. The design of this study and the resulting design principles will combine to answer the following research question:

How can a complex adaptive systems conceptualisation help meso-level practitioners enhance and transform university learning analytics implementation capability?

3 AN OUTLINE OF THE CURRENT KNOWLEDGE OF THE PROBLEM DOMAIN AND THE STATE OF EXISTING SOLUTIONS

Much of the existing knowledge around learning analytics implementation is conceptual and presents idealized models of learning analytics (Colvin et al., 2017). The embryonic nature of learning analytics as a discipline and the complexity of issues that influence to its systemic uptake go some way to explaining the paucity in large-scale implementations and highlights a need for further empirical and methodological studies (Colvin et al., 2017). Although learning analytics implementation related issues are receiving some attention in the research (Ferguson et al., 2014; Macfadyen et al., 2014), there remains little guidance for institutions pursuing their own implementations. Despite the significant investments that have been made in learning analytics around the world, there are few large-scale implementations that are having demonstrable impact on learning and teaching (Colvin et al., 2017). In addition to increasingly sophisticated data capture and analysis mechanisms, learning analytics requires interpretation and meaning-making, both of which are contingent upon a sound understanding of the specific institutional context (Macfadyen & Dawson, 2012). This, coupled with an emerging concept like learning analytics, limits the usefulness of commercial-off-the-shelf and one-size-fits-all solutions (Dede, 2008).

4 A DISCUSSION OF HOW THE DOCTORAL PROJECT'S SUGGESTED SOLUTION IS DIFFERENT, NEW, OR BETTER THAN EXISTING APPROACHES TO THE PROBLEM

The orthodox approach with emerging technology related concepts in universities is to apply deliberate strategy and detailed planning (Kezar, 2001; Reid, 2009). This paradigm is well established, despite its potentially negative impact on learning analytics implementation. The application of a complex adaptive systems lens to the problem of learning analytics implementation affords an alternative approach within this strategic, plan-driven environment. Complex adaptive systems theory acknowledges the symbiotic relationship that agents have with their environments (Holland, 2014) and provides a way forward where adaptive challenges conflict with the strategic environment (Uhl-Bien et al., 2007). The use of complexity theory as a way forward with learning analytics adoption has been suggested by a number of researchers and practitioners (Colvin et al., 2017; Tsai & Gasevic, 2016). The use of a complex adaptive systems lens allows us to look beyond mechanical and hierarchical approaches to learning analytics implementation, and is one of the key differences with the approach proposed by this study.

Another important consideration in these strategic environments is the role of meso-level practitioners. Much of the work of implementation with learning and teaching related technologies occurs at the meso-level, the intermediate level between small scale, local interactions and large scale institutional strategies (Hannon, 2013). The lack of attention that is given to what actually happens in practice has been identified as a factor contributing to high failure rates of information technology implementations (Beer, Jones, &

Clark, 2012; Goldfinch, 2007; Hannon, 2013). It is proposed that the meso-level practitioners will provide an interpretive layer between strategy and planning required at the macro-level, and the learning and improvisation required at the micro-level. From a learning analytics perspective, a focus on this level can help universities design and implement purposeful and adaptable learning analytics through an iterative and consultative process. The meso-level focus and complex adaptive systems conceptualisation proposed by this study represent a new method for solving an increasingly important problem.

5 A SKETCH OF THE RESEARCH METHODOLOGY AND IDENTIFICATION OF CORE METHODS/TECHNIQUES

Understanding how a complex adaptive systems approach can contribute to institutional learning analytics implementations requires a methodological approach that complements the situated, distributed and dynamic complexity of the socio-technical systems involved. Design based research (DBR) is an applied and pragmatic methodology that aligns with the complex adaptive systems theoretical perspective of this study. DBR acknowledges that contexts, including physical, social, cognitive and cultural elements, cannot be separated from the system (Larsen-Freeman & Cameron, 2008). DBR is situated in real-world contexts and is centred around a practical problem or issue (Herrington, McKenney, Reeves, & Oliver, 2007; McKenney & Reeves, 2013). DBR allows for the development of design principles that can be used to guide learning analytics implementations. The design principles will be developed at a theoretically and pragmatically relevant research site (Ritchie, Lewis, Nicholls, & Ormston, 2013). CQUniversity was selected as the site for this study. Selection was based on criteria of suitability and pragmatism, and this university has a current issue linked with learning analytics.

This researcher works for CQUniversity and has co-developed a number of widely-used learning analytics tools within the organisation. Knowledge of, and experience with, the specific context is crucial when dealing with complex adaptive systems (Davis & Sumara, 2007). A significant gap has been identified at CQUniversity relating to students at risk of failure and/or attrition. An analysis has identified that students who are struggling in individual units are most often struggling across multiple units in that term and are consequently at greater risk of dropping out. The existing learning analytics tools were designed to afford intervention at the individual unit level. Students struggling across multiple units, students at-risk of attrition, are beyond the scope of the current suite of tools. Attrition research in the CQUniversity context has shown that students struggle with their studies due to a complex entanglement of factors, most of which are beyond the university's current ability to perceive, much less control (Beer & Lawson, 2016, 2017). This represents a complex challenge for learning analytics. The first phase of DBR constitutes the analysis of this learning analytics gap in collaboration with practitioners, other researchers and experts from both learning analytics and student retention/attrition backgrounds. As a starting point for this analysis, an initial set of design principles based on the role of the meso-level practitioners and the identified gap, have been developed for review and consultation. The opportunity to discuss these initial principles and the practical problem at LAK using semi-structured interviews will further refine the principles to be used for learning analytics design and implementation in the second DBR phase. The refined principles will guide the development, refinement and operation of a learning analytics intervention in alignment with the DBR cycles.

6 CURRENT STATUS OF THE WORK AND RESULTS ACHIEVED SO FAR

The researcher has recently successfully completed his PhD Confirmation of Candidature and is now further developing refining the design principles, planning the intervention and its associated cycles of evaluation. A decline in student retention and an increase in non-completion of units at the research site has created additional strategic impetus for projects that can help with student attrition. A number of qualification teams at the research site have already expressed their interest in participating in this research, as they struggle to holistically monitor their students' engagement amidst a cohort whose lives are becoming increasingly busy and complex. Participation in the LAK doctoral consortium will provide valuable feedback on how learning analytics can be employed to solve a complex and entangled problem.

7 REFERENCES

Beer, C., Jones, D. T., & Clark, D. (2012). Analytics and complexity: Learning and leading for the future. Paper presented at the Proceedings of the 29th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE 2012).

Beer, C., & Lawson, C. (2016). The problem of student attrition in higher education: An alternative perspective. Journal of Further and Higher Education, 1-12. doi:10.1080/0309877X.2016.1177171

Beer, C., & Lawson, C. (2017). Framing attrition in higher education: A complex problem. Journal of Further and Higher Education, 1-12. doi:10.1080/0309877X.2017.1301402

Boustani, M. A., Munger, S., Gulati, R., Vogel, M., Beck, R. A., & Callahan, C. M. (2010). Selecting a change and evaluating its impact on the performance of a complex adaptive health care delivery system. Clinical Interventions In Aging, 5, 141-148.

Clow, D. (2014). Data wranglers: Human interpreters to help close the feedback loop. Paper presented at the Proceedings of the fourth international conference on learning analytics and knowledge, Indianapolis, IN, USA.

Colvin, C., Dawson, S., Wade, A., & Gasevic, D. (2017). Addressing the Challenges of Institutional Adoption. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), Handbook of Learning Analytics (Vol. 1, pp. 281 - 289). Australia: Society for Learning Analytics Research.

Davis, B., & Sumara, D. (2007). Complexity science and education: Reconceptualizing the teacher's role in learning. Interchange: A Quarterly Review of Education, 38, 53-67.

Dede, C. (2008). Theoretical perspectives influencing the use of information technology in teaching and learning International Handbook of Information Technology in Primary and Secondary Education (pp. 43-62). Boston, MA: Springer.

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). Setting learning analytics in context: overcoming the barriers to large-scale adoption. Paper presented at the Fourth International Conference on Learning Analytics And Knowledge, Indianapolis, USA. doi: 10.1145/2567574.2567592

Goldfinch, S. (2007). Pessimism, computer failure, and information systems development in the public sector. Public Administration Review, 67(5), 917-929.

Hannon, J. (2013). Incommensurate practices: Sociomaterial entanglements of learning technology implementation. Journal of Computer Assisted Learning, 29, 168-178.

Herrington, J., McKenney, S., Reeves, T., & Oliver, R. (2007). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2007, Chesapeake, VA, USA.

Holland, J. H. (2006). Studying complex adaptive systems. Journal of Systems Science and Complexity, 19, 1-8. doi:10.1007/s11424-006-0001-z

Holland, J. H. (2014). Complexity: A very short introduction. Oxford, England: Oxford University Press.

Jones, C., Dirckinck ‐ Holmfeld, L., & Lindström, B. (2006). A relational, indirect, meso-level approach to CSCL design in the next decade. International Journal of Computer-Supported Collaborative Learning, 1(1), 35-56.

Jones, D. T., & Clark, D. (2014). Breaking Bad to Bridge the Reality / Rhetoric Chasm. Paper presented at the ASCILITE2014 Rhetoric and Reality, Dunedin, New Zealand. Conference publication retrieved from http://ascilite2014.otago.ac.nz/

Juarrero, A. (1999). Dynamics in action: Intentional behavior as a complex system. Cambridge, Massachusetts: MIT press.

Kezar, A. (2001). Understanding and facilitating organizational change in the 21st century. ASHE-ERIC higher education report, 28(4), 147.

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. IBM Systems Journal, 42, 462-483. doi:10.1147/sj.423.0462

Larsen-Freeman, D., & Cameron, L. (2008). Research methodology on language development from a complex systems perspective. The Modern Language Journal, 200-213.

Macfadyen, L. P., & Dawson, S. (2012). Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. Journal of Educational Technology & Society, 15(3), 149-163.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gasevic, D. (2014). Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Challenge. Research & Practice in Assessment, 9(Winter, 2014), 11.

McKenney, S., & Reeves, T. C. (2013). Systematic review of design-based research progress is a little knowledge a dangerous thing? Educational Researcher, 42(2), 97-100.

Motz, B. A., Teague, J. A., & Shepard, L. L. (2015). Know Thy Students: Providing Aggregate Student Data to Instructors.

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. BMJ: British Medical Journal, 323, 625-628.

Reid, I. C. (2009). The contradictory managerialism of university quality assurance. Journal of Education Policy, 24(5), 575-593. doi:10.1080/02680930903131242

Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (2013). Qualitative research practice: A guide for social science students and researchers. London, England: Sage.

Siemens, G., & Baker, R. (2012). Learning analytics and educational data mining: towards communication and collaboration. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada. http://delivery.acm.org/10.1145/2340000/2330661/p252-siemens.pdf?ip=138.77.2.133&acc=ACTIVE%20SERVICE&CFID=145100934&CFTOKEN=10069569&__acm__=1345679404_18f65a315d7b4ba9014a8f150ad6189c

Tsai, Y.-S., & Gasevic, D. (2016). Adoption of Learning Analytics in Higher Education: State, Challenges, and Policies. Retrieved from http://sheilaproject.eu/wp-content/uploads/2016/06/Adoption-of-Learning-Analytics-in-Higher-Education_Executive-Summary.pdf

Uhl-Bien, M., Marion, R., & McKelvey, B. (2007). Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. The Leadership Quarterly, 18, 298-318. doi:10.1016/j.leaqua.2007.04.002

Developing a Learning Analytics Adoption Framework for K-12 Schools

Justin T. Dellinger
University of Texas at Arlington
jdelling@uta.edu

ABSTRACT: Much of the research in learning analytics has focused on post-secondary institutions and massive open online courses, and K-12 schools have not received the same attention. Additionally, as the field of learning analytics continues to mature, researchers have increasingly investigated meta-issues such as ethics, policy, and adoption. Over the past three years, projects have studied current practices and challenges around the adoption and implementation of learning analytics at universities, but no significant effort in the K-12 space has taken place. The aim of my doctoral research is to identify essential factors that serve as barriers to or facilitate adoption in this space. To do so, I will complete a systematic literature review of research and posted policies, and interview senior leaders in school districts. After analyzing the collected data, I will synthesize the findings into a preliminary framework and evaluate it by surveying a range of school leaders.

Keywords: Learning analytics, adoption, policy, K-12, schools, challenges

1 BACKGROUND

Learning analytics has rapidly developed as an emerging field of research and practice and continues to gain attention from educators worldwide given its potential to provide data-driven solutions to improve student outcomes. While the field holds great promise to promote positive change, most of the research has occurred in postsecondary institutions and massive open online courses (MOOCs). As the field continues to mature, more emphasis is needed on the use of learning analytics in a K-12 schools context. As observed in the LAK18 conference webpage, the organizers have recognized the gap and have explicitly called for increased participation in the event (LAK18 General Call, 2017).

Another key development in the field of learning analytics is growing investigation into the adoption and implementation of learning analytics. The LAK18 conference organizers have also recognized this trend, with specific focus on the engagement of stakeholders in the design, deployment, and assessment of learning analytics (LAK18 General Call, 2017). Following the larger trends of the field, research on learning analytics adoption has also primarily taken place for higher education. My research will focus on this gap and explore the adoption of learning analytics in a K-12 context.

A number of school districts have begun to adopt and implement learning analytics. As mentioned in the most recent K-12 Horizon Report (2017), school leaders are eager to use it to guide decision-making and empower teachers through learning data to improve instruction and student outcomes. Additionally, the authors identified the use of learning analytics as a key three-to-five-year trend. However, without sufficient empirical evidence to inform adoption, K-12 school districts may not effectively and efficiently implement learning analytics.

In particular, school superintendents have significant potential to be change agents in their school districts, especially in the area of technology adoption (Ash, 2014; Maxwell, Locke, & Scheurich, 2013; Sherman, 2007). Ash notes, however, that "few of them know how to do it" (p. 4). Additionally, schools have continued falling into the trap of techno-solutionism, where educational leaders attempt to solve broad, complex problems using a narrow technological solution (Hardesty Bray, 2007; Morozov, 2013), but schools that most need the benefits that innovative technologies can offer are often highly resistant to changing the status quo and the last to adopt them (Rogers, 2003). Educational technology may also simultaneously hold great potential for some, but further social inequity for others if poorly implemented (Tawfik, Reeves, & Stich, 2016).

Increased public accountability centered on educational budgets plays a significant factor in the continued use of technology in school districts, and mismanagement by administrators can lead to trouble with state and federal policy makers (Ross, 2015). In a broader context, technology waste due to poor implementation may have a devastating effect on schools that already have scarce resources, which can create further resistance to needed changes, and teachers often do not have the necessary support to effectively implement technology in their practice (Mohammed & Harlech-Jones, 2008; Norris, Sullivan, Poirot, & Soloway, 2003). Therefore, it is increasingly important to provide well-researched frameworks that can inform the adoption process of emerging tools such as learning analytics.

2 GOALS AND RESEARCH QUESTIONS

The primary goal of my doctoral research is to investigate key factors around the adoption and implementation of learning analytics and create an evidence-based policy framework that K-12 administrators can use in their school districts. My preliminary overarching research questions include:

1) Why do school leaders adopt learning analytics?

2) Do school leaders use a specific policy framework when adopting learning analytics?

3) What challenges do school leaders face when adopting learning analytics?

4) How aware are school leaders of the quality of third-party learning analytics tools?

5) Do school leaders include stakeholders in the adoption process and in what ways?

3 CURRENT KNOWLEDGE

While there is a greater amount research around ethics, privacy, and codes of practice (Pardo & Siemens, 2014; Prinsloo & Slade, 2017; Rubel & Jones, 2016; Sclater, Peasgood, & Mullan, 2016), learning analytics adoption is a relatively new topic of inquiry. Each of these meta-issues (Ochoa & Wise, 2017) are intertwined to varying degrees, but adoption has the greatest dearth of established research. To date, there have only been a few major systematic studies about learning analytics adoption and implementation in Australia and Europe. However, researchers situated these studies in a postsecondary context and there is little established work for K-12 schools.

Dawson et al. (2017) undertook a project that examined learning analytics adoption in higher education institutions across Australia. They found that broad implementation remains at a very early stage and very few institutions were sophisticated enough to link learner performance data with practice in order to improve the learning process and student outcomes. The researchers' aim was investigate the progress of implementations to determine the state of the field and understand barriers and challenges to adoption. Overall, they found that successful adoption necessitates quick recognition and responses to organizational culture and stakeholders (Dawson, 2015).

Similarly, researchers for the SHEILA project have examined the progress of learning analytics implementation to help European universities transition in an increasingly digital world (SHEILA About Page, n.d.). As with the study in Australia, most European higher education institutions are at very early stage with regard to the adoption of learning analytics. The researchers aim to complete a policy development framework that can address the gap around the use of student data to inform policy and practice.

Halliday and Anderson (2016) have also explored the creation of a framework to help higher education institutions in the United Kingdom better use and visualize learning analytics. As with the Australia and SHEILA studies, the researchers found no standardized effort to adopt and implement learning analytics. Additionally, they found that there was still not enough evidence in their study around the impact of learning analytics on behaviors and student outcomes, even though they had some promising early results.

As part of my article-based dissertation, I plan to author a conceptual paper on the potential challenges for the adoption and implementation of learning analytics in K-12 schools by conducting a review of current literature and searching school district websites for any posted policies about this topic. I will use the work by Tsai and Gasevic (2017) as the basis of my investigation. In their study, they used keyword searches such as 'learning analytics' in journals and databases to identify appropriate sources. Tsai and Gasevic also used a snowball approach by mining citations and references for additional sources. They then filtered themes by specific criteria and incorporated similar topics if they met that criteria (e.g., including 'academic analytics' if it was applicable in a learning analytics context). By doing so, they were able to describe the state of learning analytics adoption in higher education and six challenges for strategic planning and policy. These include a shortage of leadership capabilities, unequal engagement with stakeholders, a lack of pedagogy-based approaches, insufficient training, few studies that empirically validate impact, and limited learning analytics specific policies. Finally, the authors identified and examined eight existing policies that include topics such as strategy, obligations, privacy, and data management/governance. Tsai and Gasevic conclude that established policies do not do enough to address learning analytics specific policies, communication between stakeholders, pedagogy-based approaches, skill development, and evaluation of effectiveness. In my third chapter, I also plan to use additional SHEILA work (Tsai et al., 2017) for the development and analysis of interviews with senior school leadership.

4 JUSTIFICATION

Since there has been no systematic work done on K-12 learning analytics adoption, the creation of an evidence-based framework could provide leaders with an effective tool to guide implementation in their schools. It is unclear if developing higher education frameworks are applicable in a K-12 context, and further research between the systems is needed.

5 PRELIMINARY RESEARCH METHODS

My dissertation will have three separate studies with one overarching theme to develop a policy framework for learning analytics adoption in K-12 schools. The introductory chapter will lay the groundwork for the three articles through discussion of the larger problem, background, purpose, significance, design, limitations, and definitions in my dissertation. The final chapter will serve as discussion of my findings and presentation of the policy framework.

As mentioned in Section 3, my second chapter will be a conceptual paper of potential challenges for the adoption and implementation of learning analytics in K-12 schools. Given that most of the work on learning analytics adoption has taken place in a postsecondary context (Dawson, 2015; Tsai & Gasevic, 2017), I plan to use the results of the SHEILA project framework as a starting place and identify overlap and gaps between K-12 schools and postsecondary institutions. There are unique differences between these two groups in a United States context, given that they developed separately for most of the 20th century (Callan et al., 2009), and it is unclear if higher education learning analytics policies are at all applicable at the K-12 level.

In my third chapter, I will use the results of the conceptual paper to inform semi-structured interviews to investigate K-12 superintendents' perceptions, motivations, and understanding about the role of learning analytics in their school districts. I will conduct a multi-site case study and interview up to six North Texas public school superintendents that are interested in or are beginning to adopt and implement learning analytics in their school districts. The methods used in Tsai et al. (2017) will serve as the basis for analysis in this chapter, as well as my fourth chapter.

At present, the plan for my fourth chapter is the least developed. This is partially because the middle chapters of my dissertation are an iterative process that build on each other. My conceptual paper will inform my interviews and analysis of the interviews and literature will lead to the creation of a preliminary framework. At that point, I intend to complete a quantitative study using surveys to test and refine key elements of the framework. These surveys will target school leaders, including individual campus administration and district policy makers.

6 CURRENT WORK STATUS

At the current stage of my doctoral program, I have not conducted dissertation-specific research, but I have aligned and focused my course papers on this topic. I finished one conceptual paper in spring 2017 on learning analytics implementation ethics and am working to submit it for publication by the end of the spring 2018 semester. In this paper, I completed a review of relevant literature focusing on ethics, privacy, and moral practice for big data and learning analytics, and how they relate to broader societal issues facing education (e.g., black box algorithms perpetuating problems of social inequality). Additionally, I am using my current qualitative methods course to develop my research design for the aforementioned interviews with K-12 school superintendents. This paper will serve as the third chapter of my dissertation.

REFERENCES

Ash, K. (2014). K-12 leaders shift from protectors of status quo to change agents. *Education Week*, 33(21), 4–5.

Callan, P. M., Kirst, M. W., Shulock, N. B., Spence, D. S., Walsh, E. J., & Usdan, M. D. (2009). The historical context of the divide between K–12 and higher education. In *States, schools, and colleges: Policies to improve student readiness for college and strengthen coordination between schools and colleges. National Center Report #09-2*. National Center for Public Policy and Higher Education. Retrieved from https://eric.ed.gov/?id=ED508093

Dawson, S. (2015). *Benchmarking learning analytics in Australia* [Slideshare Slides]. Retrieved from https://www.slideshare.net/shaned07/benchmarking-learning-analytics-in-australia

Freeman, A., Adams Becker, S., Cummins, M., Davis, A., & Hall Glesinger, C. (2017). *NMC/CoSN Horizon Report: 2017 K-12 edition*. Austin, Texas: The New Media Consortium. Retrieved from https://www.nmc.org/publication/nmccosn-horizon-report-2017-k-12-edition/

Halliday, J., & Anderson, M. (2016). Developing a framework for the visualisation of learning analytics in UK higher education. In M. Anderson & C. Gavan (Eds.), *Developing effective educational experiences through learning analytics* (pp. 119–142). Hershey, PA: IGI Global.

Hardesty Bray, S. (2007). E-Learning: Successes and failures. *Chronicle of Higher Education*, *53*(18), B20–B23.

LAK18 General Call. (2017). Retrieved from https://latte-analytics.sydney.edu.au/general-call/

Learning Analytics in Australia Home Page. (2017). Retrieved from http://he-analytics.com/

Maxwell, G. M., Locke, L. A., & Scheurich, J. J. (2013). Case study of three rural Texas superintendents as equity oriented change agents. *The Qualitative Report*, *18*(11), 1–23.

Mohammed, R. F., & Harlech-Jones, B. (2008). The fault is in ourselves: Looking at “failures in implementation.” *Compare: A Journal of Comparative Education*, *38*(1), 39–51. https://doi.org/10.1080/03057920701420825

Morozov, E. (2013, March 2). *The perils of perfection*. The New York Times. Retrieved from http://www.nytimes.com/2013/03/03/opinion/sunday/the-perils-of-perfection.html

Norris, C., Sullivan, T., Poirot, J., & Soloway, E. (2003). No access, no use, no impact: Snapshot surveys of educational technology in K-12. *Journal of Research on Technology in Education*, *36*(1), 15–27.

Ochoa, X., & Wise, A. (2017, March). *2017 Learning Analytics & Knowledge Conference Opening Remarks*. Slides presented at the 7th International Learning Analytics & Knowledge Conference, Simon Frasier University. Retrieved from https://www.slideshare.net/alywise/lak17-learning-analytics-conference-opening-remarks

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438–450.

Prinsloo, P., & Slade, S. (2017). Big data, higher education and learning analytics: Beyond justice, towards an ethics of care. In B. K. Daniel (Ed.), *Big data and learning analytics in higher education* (pp. 109–124). Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_8

Rogers, E. M. (2003). *Diffusion of innovations*. New York, NY: Free Press.

Rubel, A., & Jones, K. M. L. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society, 32*(2), 143–159. https://doi.org/10.1080/01972243.2016.1130502

Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education*. Jisc. Retrieved from https://www.jisc.ac.uk/reports/learning-analytics-in-higher-education

SHEILA About Page. (n.d.). Retrieved from http://sheilaproject.eu/about-sheila/

Sherman, W. H. (2007). No child left behind: A legislative catalyst for superintendent action to eliminate test-score gaps? *Educational Policy*, *22*(5), 675–704. https://doi.org/10.1177/0895904807307063

Tawfik, A., Reeves, T., & Stich, A. (2016). Intended and unintended consequences of educational technology on social inequality. *TechTrends: Linking Research & Practice to Improve Learning*, *60*(6), 598–605. https://doi.org/10.1007/s11528-016-0109-5

Tsai, Y.-S., & Gasevic, D. (2017). Learning analytics in higher education — Challenges and policies: A review of eight learning analytics policies. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 233–242). New York, NY: ACM. https://doi.org/10.1145/3027385.3027400

Tsai, Y.-S., Gasevic, D., Muñoz-Merino, P. J., & Dawson, S. (2017). La policy: Developing an institutional policy for learning analytics using the RAPID outcome mapping approach. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 494–495). New York, NY: ACM. https://doi.org/10.1145/3027385.3029424

Learners' Self-Regulated Learning Strategies in Massive Open Online Courses: Metrics and Variables that Influence Performance

Jorge Maldonado-Mahauad
Departamento de Ciencias de la Computación
Pontificia Universidad Católica de Chile, Chile
Universidad de Cuenca, Ecuador
jjmaldonado@uc.cl

ABSTRACT: Massive Open Online Courses (MOOC) have become a source of digital content that can be tackled timelessly and from anywhere. These offer quality content to millions of learners around the world, providing new opportunities for learning. However, only a fraction of those who initiate a MOOC can complete it, leaving thousands of committed students without achieving their goals. Recent research suggests that the main reason why students find it difficult to complete these courses is because they have problems planning, executing, and monitoring their learning process autonomously; that is, they do not effectively self-regulate their learning (SRL) in these types of environments. The objective of this PhD is to define metrics to measure and identify SRL strategies that learners employ in MOOCs and that relate with academic achievement and investigate what are the variables of the learner, the MOOC and the context in which the course is used, that influence in the use of these strategies.

Keywords: Massive Open Online Courses, Self-regulation, Learning, Metrics, Instruments.

1 CONTEXT OF THE RESEARCH PROBLEM

1.1 Massive Open Online Courses

MOOCs offer open (and in most cases free-of-charge) high quality online education to an unlimited number of participants, regardless of their origin or level of education. Thanks to MOOCs, learners can now independently access and explore enormous amounts of relevant content. Nevertheless, studies based on real data indicate that two phenomena belie this idea (Jaggars, 2014; Kay, 2013): 1) only few students complete these courses; and 2) most of the learners enrolling and completing MOOCs are highly educated people who hold some previous academic degree (Nesterko et al., 2014; Ho et al., 2014).

Recent research suggests that students experience difficulties in planning, implementing and monitoring their learning process in these types of courses autonomously; that is, they do not effectively self-regulate their learning to successfully complete a MOOC (Laplante, 2013; Kay et al., 2013). In MOOCs, learners' success relies heavily on their ability to independently and actively engage in the learning process (Weller 2011; Wang et al, 2013). This research also shows that students in a MOOC have difficulty with the effective use of SRL strategies. These strategies are used to self-regulate cognitive level (process volumes of content), meta-cognitive (planning and organization to achieve the proposed objectives) and resource management (time management, seeking help, adapting the study environment) (Kizilcec et al., 2015). In this milieu, Self-regulated

Learning (SRL) strategies play a key role in success. But, how to support learners with SRL strategies in MOOCs and what the associations between strategies and achievement are, still remain open challenges. ***Problem 1: MOOC learners have difficulties finishing the course because they lack (meta)cognitive and resource management SRL strategies required for organizing their learning and environment.***

1.2 Self-regulated Learning in MOOCs: strategies and variables

Self-regulation of learning has been studied on the basis of several self-regulated learning models, based on different theories offering definitions and concepts that allow us to understand it (Panadero, 2017). However, all models share common assumptions, which allow us to understand SRL as an active construction process through which students define objectives for their learning and attempt to monitor, regulate and control their cognition, intentions and behaviour, guided and restricted by the objectives and contextual characteristics of the environment (Pintrich, 2000). The Pintrich's model of SRL (1999) identifies three categories of SRL strategies that students should use to regulate their own learning: 1) Cognitive strategies that relate to the activities students use in the acquisition, storage and retrieval of information; 2) Meta-cognitive strategies that relate to the activities undertaken by students to monitor themselves and monitor learning, and 3) Resource management strategies that refer to the activities students undertake to manage the learning environment and resources provided.

However, Jakesová & Kalenda (2015) point out that one of the main difficulties when researching SRL is that the existing models for analyzing how SRL occurs are not easy to operationalize. According to the authors, an alternative direction for scientific research in studying SRL should be based on the search for "specific causal mechanisms instead of great theories describing results of self-regulated learning with no regard to the variability of its context, contents and participants," (Jakesová and Kalenda, 2015). In MOOCs, where there is data available from different learners (participants) and interactions (events) with a MOOC-based environment (context and contents), it is possible to understand causal relationships based on actual data. In this sense, recent research on MOOCs suggests that there are characteristics of context, MOOC and participants that may influence how a student's SRL strategies are developed in these courses. This can be used as a basis for understanding which variables should be considered to analyze student SRL strategies in a MOOC. ***Problem 2: Determining SRL strategies of learners in MOOCs, and the variables influencing these strategies are still open challenges. Current models and frameworks of SRL are not appropriate for analyzing and describing the SRL strategies employed by learners in MOOCs. New frameworks of analysis are required in order to understand how MOOC learners address SRL strategies and what those contextual, course content, and personal variables that influence them are.***

1.3 Measuring effectiveness of SRL strategies

Learners' SRL strategies have been traditionally evaluated in terms of achievement (Taub et al, 2014). In MOOCs, up until now, achievement has been related mainly to course completion. To understand why learners complete a MOOC or not, current studies have focused on analyzing students' engagement and disengagement patterns with course content such as lectures, assessment activities, discussion forums and social media complementing MOOCs (Clow, 2013; Guo et al, 2014; Li et al, 2015; Joksimović et al, 2015). These patterns have been used to categorize learners in order to make informed future interventions in the course design that address their needs (Kizilcec et al, 2013; Ferguson & Clow 2015; Kizilcec & Schneider, 2015). However, these patterns do not give insights into how learners reached each status, or why those completing the course actually finished it (Zheng et al, 2015). Recent studies indicate that completion of a course is not necessarily the best measure for achievement in MOOCs (Kizilcec, 2015).

Depending on the model that is assumed, self-regulation in a MOOC can be studied from two perspectives: (1) as an aptitude or (2) as an event (Wirth & Leutner, 2008). In the case of the first perspective, many instruments have been developed in the last decade to measure the level of use of SRL in online environments, being one of the most commonly used the questionnaires (Roth et al., 2015). In the case of the second perspective the study of self-regulation as an event has gained strength, since it is through the traces of data left by students as a result of interaction with the course, that part of the students' cognitive activity can be evidenced while self-regulating their learning process (Jakešová & Kalenda, 2015). In this sense, thanks to the advance of computational techniques such as process mining, it is possible to describe the models of learning processes that can represent learning activities through petri networks or flowcharts providing new robust ways to extract, analyze and visualize the traces of student activities (Mukala, Buijs, Leemans, & van der Aalst, 2015). ***Problem 3: The concept of academic achievement or completion in a MOOC must be redefined from the student's perspective (learning objective). New analyses are needed to understand the relationship between SRL strategies that students report on themselves, the strategies they actually use, and their relation with academic achievement.***

2 PROPOSAL DESCRIPTION

This research project aims to define metrics to measure and identify SRL strategies that learners employ in MOOCs and that relate with academic achievement and investigate what are the variables of the learner, the MOOC and the context in which the course is used, that influence in the use of these strategies.

2.1 Analysis of SRL strategies in MOOCs: variables and measurements

The first goal is to define metrics to measure and analyze learners' SRL strategies in MOOCs. These metrics will be based on the literature and will serve as a basis for the analysis of MOOC variables that influence the effective use of self-regulatory strategies that students employ in this type of course. To this end, the different MOOCs deployed in different contexts (on-demand and hybrid) and the main variables to be considered will be analyzed. For the definition of this framework, we adopt the perspective set by Jakešová and Kalenda (2015) and consider not only current SRL theories, but also the actual behaviors of learners in the course. As we showed in section 1.2, literature on both MOOCs and SRL evidences that context, content and participants' variables have an effect on how learners interacts and addresses a MOOC.

Accordingly, we propose in this project a preliminary framework that considers these variables organized into three groups: Learners' personal variables, context and MOOC variables (Fig. 1). Learners' personal variables refer to the attributes of the participant (demography, intentions, and prior knowledge), and his/her behavior in the course (interactions with course content). MOOC variables refer to the interplay between the course content (type of resources, structure and nature of the tasks – collaborative or individual), and context in which the MOOC is applied (Online). The following research question is derived: **Research Question 1.** *What variables from the learner, context and MOOC influence in the effective use of SRL strategies that students employ in these types of courses?*

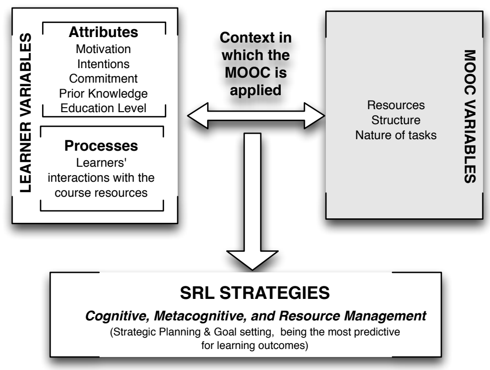


Figure 1: Preliminary framework of variables that influence MOOC students' SRL strategies.

2.2 Metrics to evaluate learners' SRL strategies in relation to achievement

The second goal is to develop tools and metrics to measure the relationship between students' SRL strategies (cognitive, meta-cognitive and resource management) and academic achievement in different types of MOOCs deployed in online contexts. For this purpose, instruments will be proposed that allow the study of self-regulation as an aptitude (self-reported) and as an event (process mining). As an aptitude and as suggested by the literature (Kizilcec & Schneider, 2015), the OLEI scale together with a questionnaire to self-report self-regulation strategies (Niemi, Nevgi, & Virtanen, 2003) will lay the foundations for creating tools to measure students' initial objectives, preferences and strategies used. As an event (process), the data resulting from student interactions with the platforms will help to identify not only what self-regulatory strategies are used to achieve academic outcomes, but also a new categorization of students according to their self-regulatory profile and performance. The following research question is derived: **Research Question 2.** *What metrics are most appropriate for understanding what self-regulatory strategies students employ in a MOOC and what effect these strategies have on their performance?*

3 RESEARCH METHODOLOGY

In this project, a mixed (qualitative and quantitative) and iterative methodology will be used as the baseline for the organization and planning. This methodology combines empirical educational research with the theoretical design of instructional strategies and tools. The phases of this methodology are: analysis, design, development/implementation and evaluation. Each of these phases will be applied in an iterative way for each of the specific goals of the research project, specifying the tasks to be developed, the methods, the data collection and analysis techniques used. Also, a literature review will be considered as an important activity during this project and will be performed for each goal. **For the first goal**, we have considered the next list of tasks: ***1) Prepare experimental scenario:*** developing MOOCs under different platforms, ***2) Define and select instruments, data collection and analysis techniques:*** the data from MOOCs will be gathered using qualitative (for deeper understanding) and quantitative techniques (showing trends). Instruments for data collection will consist of: 1) Instruments for repeated measurements (e. g. self-report questionnaires with framework analysis variables); 2) Instruments for process-oriented data processing (e. g. MOOC log files); and 3) Instruments for measuring student outcomes (e. g. final grades), ***3) Data collection:*** the data collected for initial analysis will be performed at three different times corresponding to different dates depending on MOOCs, ***4) Data analysis and preliminary framework review:*** the quantitative data collected by the questionnaires and log files on step 3 will be statistically analyzed to understand the type of interactions that students in MOOC have used in relation to the learning strategies they indicate. **For the second goal**, we have considered the next list of tasks: **1) Prepare data for analysis**: extract and clean data, **2) Prepare analysis tools**: developing scripts for analyse data gathered, **3) Data analysis:** Definition of academic achievement and student categorization: use of combined techniques from data mining (e.g. clusterization) and process mining (e.g. model discovery), **4) Determine the relationships between academic achievement (performance) and self-regulatory strategies:** based on the results of the step 3 we will use regression models to predict performance based on learners characteristics and learning sequence patterns extracted from the analysis and relate with SRL strategies in the literature.

4 STATUS OF THE WORK AND CONTRIBUTIONS

This PhD project started on August 10 of 2015 under the direction of the professor Mar Pérez-Sanagustín at the Pontificia Universidad Católica de Chile. The duration of the PhD program is about 4 years. I have started the third year, for this reason is very important to meet at the conference specialist in the field of Learning Analytics and present the advances of our research in order to obtain feedback and advice that could help us to advance in this research project. During the two years of research we have achieved the following results:

Related Goal 1 - Analysis of SRL strategies in MOOCs: variables and measurements: We evaluated how SRL strategies were related to achieving three different personal course goals: 1) earning a course certificate, 2) completing assessments (independent of grades), and 3) watching video-lectures in the course. We found that goal setting and strategic planning were significant positive predictors of goal attainment for all three goals. Also, we assessed individual differences in self-reported SRL strategies based on 27 individual characteristics, encompassing demographics, prior experience, time commitment, goals and motivations. We found that older learners reported consistently higher levels of SRL, except for help seeking. Women reported lower levels of strategic planning, elaboration, and self-evaluation. Learners with a Ph.D. reported generally strong SRL skills, they reported being much less inclined to seek help. These results were published in the Journal Computers and Education: *Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. Computers & education, 104, 18-33.*

Related Goal 2 - Metrics to evaluate learners' SRL strategies in relation to achievement: We developed and validated a questionnaire for measure SRL strategies in MOOCs. This questionnaire is composed of 22 items to assess 5 SRL strategies in MOOCs. Also, using process mining we extracted interaction sequences from fine-grained behavioural traces for 3,458 learners across three Massive Open Online Courses. As a result, we identified six distinct interaction sequence patterns. We matched each interaction sequence pattern with one or more theory-based SRL strategies and identified three clusters of learners. First: Comprehensive Learners, second: Targeting Learners, third: Sampling Learners. These results have been submitted to the Journal Computers in Human Behavior: *Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, René F. Kizilcec, Nicolás Morales, Jorge Munoz-Gama (2017). Mining Theory-Based Patterns from Big Data: Identifying Self-Regulated Learning Strategies in Massive Open Online Courses (Accepted)*

REFERENCES

Clow, D., 2013, April. MOOCs and the funnel of participation. In Proceedings of the Third International Conference on Learning Analytics and Knowledge (pp. 185-189). ACM.

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of mooc videos. Learning@ scale conference, 41-50.

Ho A. D, Chuang, I., Reich, J., Coleman, C. A., Whitehill, J., Northcutt, C. G., Williams, J. J., Hansen, J., D., Lopez, G., Petersen, R. (2015). HarvardX and MITx: Two Years of Open Online Courses Fall 2012- Summer 2014, Social Science Research Network

Jaggars, S. S. (2014). Democratization of Education for Whom? Online Learning and Educational Equity, Diversity and Democracy, 17, 1

Jakešová, J., & Kalenda, J. (2015). Self-regulated Learning: Critical-realistic Conceptualization. Procedia-Social and Behavioral Sciences, 171, 178-189.

Joksimović, S., Kovanović, V., Jovanović, J., Zouaq, A., Gašević, D., & Hatala, M. (2015, March). What do cMOOC participants talk about in social media?: a topic analysis of discourse in a cMOOC, 5th ACM International Conference on Learning Analytics And Knowledge, 156-165.

Kay, J. (2013). MOOCs: So Many Learners, So Much Potential...," IEEE IS, 28(3), 70–77.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. LACK, 170-179.

Kizilcec, R.F. and Schneider, E., 2015. Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. ACM Transactions on Computer-Human Interaction (TOCHI), 22(2), p.6.

Laplante P. A. (2013) Courses for the Masses?. IT Professional, 15(2), 57– 59.

Li, N., Kidziński, Ł., Jermann, P., & Dillenbourg, P. (2015) How Do In-video Interactions Reflect Perceived Video Difficulty? eMOOCs Conference.

Mukala, P., Buijs, J. C., Leemans, M., & van der Aalst, W. M. (2015, December). Learning Analytics on Coursera Event Data: A Process Mining Approach. In SIMPDA (pp. 18-32).

Nesterko, S. O., Seaton, D. T., Kashin, K., Han, Q., Reich, J., Waldo, J., Chuang I., Ho, A.D. (2014). Education Levels Composition (HarvardX Insights)

Niemi, H., Nevgi, A., & Virtanen, P. I. (2003). Towards self-regulation in web-based learning. Journal of Educational Media, 28(1), 49-71.

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. Frontiers in Psychology, 8.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In Handbook of self-regulation (pp. 451-502).

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge?. Computers in Human Behavior, 39, 356-367.

Weller, M. (2011). A pedagogy of abundance. Revista Española de Pedagogía, 223-235.

Wirth, J., & Leutner, D. (2008). Self-regulated learning as a competence: Implications of theoretical models for assessment methods. Journal of Psychology, 216(2), 102-110.

A learning analytics approach to scaffolding scientific modeling in the classroom

David Quigley
University of Colorado Boulder
daqu3650@colorado.edu

ABSTRACT (Doctoral Consortium): I use learning analytics to understand and support scientific modeling in the classroom. In particular, I study ways we can automatically characterize student modeling, design tools to promote successful modeling, and what impact our modeling activities have on students' engagement and learning. My work focuses on student use of EcoSurvey, a modeling tool designed to help students map the organisms and interactions in their local ecosystem. Through this work, I have developed techniques that evaluate temporal sequences of student modeling activity to determine differences in student use of modeling practices, implemented statistical measures of models to determine model strength, and developed a method to blend iteration and low-impact student surveys with analytics to determine the efficacy of my work.

Keywords: Scientific Modeling, Temporal Analytics, Student Impact

1 INTRODUCTION

My research studies students' scientific modeling. In particular, I focus on how we can automatically characterize students' scientific models and their modeling practices, how the design of digital modeling tools can support student modeling practices, and how modeling practices influence student engagement and learning. I study these questions through the iterative design and deployment of a digital modeling tool, developing and applying a framework for normalizing modeling actions across tools and analytics for characterizing modeling and practices in real time. I also explore types of feedback that are useful for teachers and students to support reflection. I aim to demonstrate that improved modeling tool design and the incorporation of real-time feedback based on novel analytics can have a positive impact on the student modeling experience.

Scientific models represent ideas, processes, and phenomena by describing important components, their characteristics, and their interactions. Constructing and using models to explain scientific phenomena is also an essential practice in contemporary science classrooms, according to *A Framework for K-12 Science Education* (NRC, 2012). However, while it is widely recognized that developing students' ability to create and use models to understand phenomena is important, learning sciences research has documented challenges to implementation in the classroom, including variations in how teachers approach modelling (Mason et. al., 2005) and variations in how students engage with the practices (Kolodner et. al., 2003). In addition, modern researchers emphasize "engagement in the science and engineering practices to develop, investigate, and use scientific knowledge" as an aspect of student learning (NRC, 2014). My research aims to address this gap in understanding scientific modeling as an aspect of three-dimensional scientific learning.

1.1 Research Context

The context for my research is scientific modeling in high school biology classrooms. My work is part of Inquiry Hub, a research-practice partnership developing inquiry-based biology curriculum for middle and high school classrooms (Leary et. al., 2016). Within the high school ecology unit, students are tasked with creating a model of their local ecosystem using EcoSurvey, a digital modeling tool designed to represent the organisms and interactions the students encounter as they map a local field site (Figure 1).

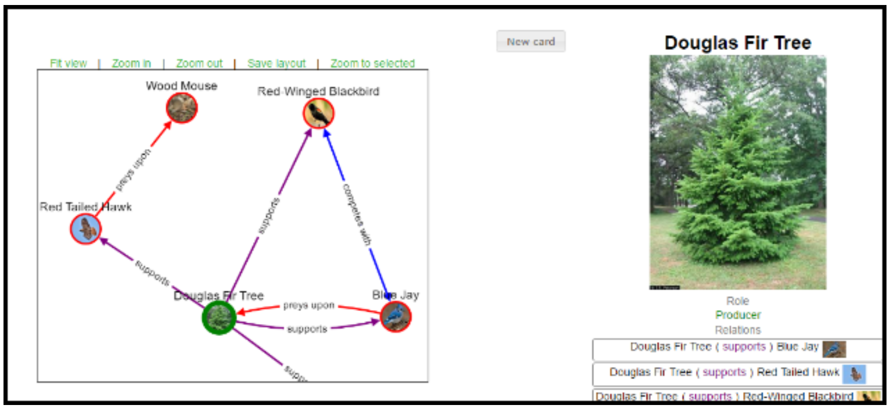


Figure 1: EcoSurvey version 2 graph view.

EcoSurvey, along with the entire Inquiry Hub biology curriculum, has undergone multiple iterations of design and deployment (Quigley et al., 2017a). Our deployments have quickly grown to incorporate approximately 1,000 students by 2016, with more growth in 2017, as seen in Table 1.

Table 1: EcoSurvey deployment measures.

Number of Teachers	Number of Classrooms	Number of Students
3	10	262
12	34	1000 (estimate)
30 (estimate)	90 (estimate)	2700 (estimate)

2 RESEARCH GOALS

2.1 RQ1: How can we automatically characterize students' models and their engagement with modeling practices at scale?

I characterize students' scientific models using a learning sciences view of modeling as a cross-cutting science concept and an important science and engineering practice (NRC, 2012). This framework drives approaches for normalizing models and the actions students take to create, revise, and use them. I study how this approach applies to the use of EcoSurvey. By normalizing models and modeling practices to existing learning sciences theory, I develop analytic approaches that examine the use of digital modeling tools from a modeling perspective, rather than from the domain-specific perspective of how these models apply to the problem at hand.

2.2 RQ2: What methods can we use to promote successful scientific modeling?

I use two different approaches to promote scientific modeling. First, I examine the design of EcoSurvey and workflow of the students' modeling process to determine how the features of our digital modeling tool map to modeling practices. Analysis of student use of EcoSurvey demonstrates the degree to which students engage with specific modeling practices. The other approach examines the students' models and modeling practices to generate customized feedback that can inform future actions. EcoSurvey incorporates systems that use an online approach to analyzing models and practices to provide support to teachers and students during the modeling process, helping teachers give strong formative feedback and helping students to understand modeling as a series of practices.

2.3 RQ3: How do students' modeling practices influence their engagement and learning?

This question characterizes the experience from the student perspective. My ideal goal is to develop a system that drives students to not only engage with modeling practices and build successful models, but also to feel engaged with the activity and context of their work. I measure this construct through experience sampling with a student self-check deployed at the end of a day's activity (Penuel et. al., 2016). I also want to support students' understanding of the material, particularly models as a cross-cutting concept.

3 RELATED WORK

My research bridges work and ideas from Learning Sciences and Computer Science to enhance students' opportunities to learn. The foundations of this work stem from scientific modeling education research, a growing discipline in the Learning Sciences space that seeks to understand the features of classroom scientific models and the processes behind students' modeling. This research demonstrates a growing body of knowledge from a cognitive perspective of student understanding. Machine Learning research provides ideas and methods for understanding these very issues; I leverage normalization and classification techniques to provide a generalized understanding of modeling without expert intervention.

3.1 Scientific Modeling Practices in Education

Understanding what students are doing at a fine-grained level can provide teachers with useful insights into learning processes, as well as provide feedback as to where and when students need additional assistance. Towards this end, several scholars have developed frameworks characterizing effective student modeling practices (e.g. Schwarz et. al., 2009). They identify a series of seven practices: (1) identifying the anchoring phenomena to be modeled, (2) constructing a model, (3) testing the model, (4) evaluating a model, (5) comparing the model against other ideas, (6) revising the model, and (7) using the model to predict or explain phenomena. Their research suggests that supporting students to engage in these practices can lead to positive learning outcomes. However, this line of work is focused on an in-depth, qualitative review of student work. In contrast, I aim to design methods that will analyze these modeling practices in an automated, scaled manner.

3.2 Online Learning Analytics

While Learning Analytics contains a strong body of existing work in understanding student activity, this work is often performed in a post-hoc fashion, extracting what improvements, differences, or gaps may exist from a tool deployment. An important area of ongoing research seeks to understand how these methods can be used to predict potential performance. Creating “online” systems (i.e. systems that measure performance and draw conclusions in real time, during use) has the potential to detect critical differences and gaps in learning and engagement. In turn, these predicted differences can be used for intervention, supporting students in useful ways. At the most recent conference on Learning Analytics & Knowledge, research teams (e.g. Bote-Lorenzo et. al., 2017) presented preliminary results in designing and/or deploying predictive systems for education and learning, while many others (e.g. Quigley et. al., 2017a, Käser et. al., 2017) cited prediction for intervention as an important next step for their research.

4 INNOVATION

My work crosses the fields of computer science, cognitive science, and education. I develop and adapt new methods of computational learning analytics to understand student activities in modeling tools, supporting an important need in science education. I examine variations in student modeling across classrooms and teachers, analyzing both students’ opportunities to learn and variations in the degree to which they engaged in specific modeling practices. I also explore the ways in which individual students’ modeling processes are indicative of student and teacher differences. Finally, I close the feedback loop by using these predictions to guide feedback within EcoSurvey.

First and foremost, my work supports teachers and students. In our preliminary work, teachers expressed a pressing need for support in analyzing and evaluating students’ modeling to provide substantive guidance through instruction, discussion, and grading. My methods create a summary view of the students’ models and activity that teachers can quickly interpret. This summary can then be leveraged in classroom practice to guide students to creating richer models.

This approach also opens new grounds in learning analytics. Normalizing models and modeling practices allows me to apply computational methods to supporting teachers in classrooms at scale. This approach allows me to generalize findings beyond the particular use case of our teachers using EcoSurvey. I also expand learning analytics to incorporate new techniques that apply to sequential activity with a digital tool.

5 METHODOLOGY

5.1 Variation in Student Scientific Models

To examine variation within student models, I evaluate the richness of students' models in terms of the number of organisms and their relationships. I also look at the balance of interactions per organism by evaluating both the average number of interactions per organism and variance in the distribution of interactions. In addition to performing richness analysis on data across deployments, I also analyze the distribution of relationship types using evenness. This measure considers how each type of relationship is represented within the survey. We calculated evenness using the shannon index, the same formula for species evenness in the study of ecosystems (Shannon et. al., 1949).

5.2 Variation in Modeling Practices

I take a dual approach of analyzing aggregate features of total activity as well as analyzing specific sequences of activity. I begin by automatically capturing usage from EcoSurvey as students use the tool to construct their digital models, creating an eight feature vector for each student consisting of total number of EcoSurvey actions, total number of create actions, total number of evaluate actions, total number of revise actions, total number of use actions, total number of EcoSurvey action types taken, number of sessions, and number of iterations. I also use action sequences to determine differences in student activity over time. In our work, a sequence pattern consists of a series of EcoSurvey actions (e.g. ”New Card” → ”Edit” → ”Generate Graph”) embedded within a student's complete action log. I treat every action as a token and determine the frequency of consecutive token sequences (n-grams), including wildcard actions (skip-grams), from these logs. To understand which sequences best characterize differences in students' modeling, I create a feature vector to represent students and use a best-first search to reduce the full set of sequence patterns to the most predictive features. I then use a Naïve Bayes classifier with variety, frequency, and iteration features and the most predictive sequence patterns to classify students according to their activity.

5.3 Influence of Tool Design on Models

My second research question seeks to understand how design changes in digital modeling tools can have an impact on students' models. To evaluate the impacts of design, I compare directly across deployments of different versions of EcoSurvey, running the same statistical comparisons for each version and compare across conditions. In cases where direct comparison of means and variance is possible, I use a t-test to determine significance. To analyze the feedback system, I plan to do direct analysis of students' models and practices as a pre/post intervention measure, with the moment of feedback as the intervention. This analysis will focus on the relative distribution of both model and practice features in each stage. The primary indicator of successful intervention will be an uptake in engagement with underutilized practices at the moment of intervention. If a student receives feedback when they have spent significant time creating and revising their model but have not revised or used their models, we should expect to see a higher frequency of revision and use actions after intervention.

5.4 Influence on Student Engagement and Learning

My analysis of the student self-check responses will focus on differences to be found based on various groupings of the responses. The first grouping of importance is the difference between responses given during EcoSurvey lessons to responses during other lessons. This comparison allows for analysis of the impact of EcoSurvey as a tool. The other comparison is to examine the possible correlation between automatic EcoSurvey engagement measures and student responses. Correlating my usage and engagement measures with student affective responses provides a great deal of evidence of the validity of my analytic methods.

6 CURRENT STATUS

So far, I have developed and deployed two iterations of the EcoSurvey tool. During the first deployment in fall 2015, we established a baseline of student models and activity as a preliminary answer to RQ1. We ran these analyses in spring 2016 and published the results at LAK 2017 (Quigley et al., 2017b). This work also informed our redesign work for the second iteration in summer 2016 and motivated the development of RQ2. The second version was deployed in fall 2016, with iterative analyses of the differences in student models between deployments accepted in the Frontiers in ICT special issue on Digital Education (Quigley et al., 2017a). My current efforts are focused on closing the analytic loop. Our third and current version of EcoSurvey incorporates the feedback mechanisms described above. We are in the middle of our deployment cycle for fall 2017, which will support my analyses during the spring and summer.

REFERENCES

Bote-Lorenzo, M.L., and Gómez-Sánchez, E. (2017). Predicting the decrease of engagement indicators in a MOOC. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (LAK '17). 143-147.

Käser, T., Hallinen, N.R., and Schwartz, D.L. (2017). Modeling exploration strategies to predict student performance within a learning environment and beyond. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (LAK '17). 31-40.

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., … Ryan, M. (2003) Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design(tm) into practice. *Journal of the Learning Sciences*, 12(4):495–547.

Leary, H., Severance, S., Penuel, W. R., Quigley, D., Sumner, T., & Devaul, H. (2016). Designing a Deeply Digital Science Curriculum: Supporting Teacher Learning and Implementation with Organizing Technologies. *Journal of Science Teacher Education*, 27(1), 61–77.

Mason, N. W. H., Mouillot, D., Lee, W. G., & Wilson, J. B. (2005). Functional Richness, Functional Evenness and Functional Divergence: The Primary Components of Functional Diversity. *Oikos*, 111(1), 112–118.

National Research Council. (2014). *Developing Assessments for the Next Generation Science Standards*. Washington, DC: The National Academies Press.

National Research Council & others. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

Penuel, W. R., Van Horne, K., Severance, S., Quigley, D., & Sumner, T. (2016). Students' responses to curricular activities as indicator of coherence in project-based science. In *Transforming Learning, Empowering Learners: Proceedings of the 10th International Conference of the Learning Sciences* (pp. 855–858). Singapore: International Society of the Learning Sciences.

Quigley, D., McNamara, C., Ostwald, J., & Sumner, T. (2017a). Using Learning Analytics to Understand Scientific Modeling in the Classroom. *Frontiers in ICT*, 1-23.

Quigley, D., Ostwald, J., & Sumner, T. (2017b). Scientific Modeling: Using learning analytics to examine student practices and classroom variation. In *The 7th International Conference on Learning Analytics & Knowledge* (LAK '17) 329–338.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., … Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632–654.

Shannon, C. E. and Weaver, W. (1998). *The mathematical theory of communication.* University of Illinois press.

The Potential of Learning Analytics to Support Peer Assessment

Kamila Misiejuk
Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway
kamila.misiejuk@uib.no

ABSTRACT: There are few studies of peer assessment in the learning analytics (LA) field. In this research the potential of LA methods to analyze peer assessment data is examined. The first study comprises an application of LA methods on a big data set coming from the use of a peer assessment tool, Peergrade. The second study analyzes the use of LA methods in order to understand how the context data, together with the Peergrade data, can provide insights into the role of peer assessment in learning and teaching.

Keywords: learning analytics, educational data, peer assessment, learning design, teacher inquiry

1. RESEARCH BACKGROUND

Learning analytics (LA) is a new emerging field that has the potential not only to support assessment for learning, but also to better understand learning processes. Ferguson et al. (2016) calls for aligning "analytics with assessment practices" as LA has the potential to change assessment practices and support "the holistic process of learning" (Ferguson et al., 2016, p. 37). *Peer assessment* is "a quantitative evaluation and qualitative feedback of a learner's performance by another learner" (Patchan et al., 2017, p. 1). Peer assessment research has a long research history of 40 years, which showed that student's feedback can be both helpful and reliable, as well as correlate high with the teacher's grading (Patchan et al., 2017; Raes et al., 2015; Li et al., 2016). Peer assessment is especially valuable in the context of MOOCs, where the ability to provide feedback by the instructor for each learner is limited (Wahid et. al, 2016). Moreover, it is also widely used at the universities (de Alfaro & Shavlovsky, 2016).

Results from Misiejuk's (2017) Master thesis, in which the proceedings from the International Conference on Educational Data Mining, the International Learning Analytics and Knowledge, and the ACM Conference on Learning at Scale were analyzed using a variety of scientometric techniques such as keyword analysis and citation analysis, revealed that the topic of peer assessment is still relatively underexplored in the field of Educational Data Sciences (EDS) (Misiejuk, 2017). Furthermore, the state-of-the-field review of learning analytics confirms this finding (Misiejuk & Wasson, forthcoming). Studies from the field of EDS focus mainly on the relationship between peer assessment and student performance (Jiang et al., 2014; Kulkarni et al., 2015; Sajjadi et al., 2016; Tritz et al., 2014; Ashenafi et al., 2016). The other significant research areas are the quality of the student grading (Xiong et al., 2010; Xiong & Litman, 2013; Raman & Joachims, 2015), the exploration of the peer assessment metrics such as accuracy, validity, and reliability (de Alfaro & Shavlovsky, 2016; Piech et al., 2013; Vogelsang & Ruppertz, 2015; Vozniuk et al., 2014), and technological improvement of the peer assessment (Kulkarni et al., 2014; Kolhe et al., 2016;

Xing et al., 2014). Even though the definition of LA includes "the measurement, collection, analysis and reporting of data about learners and their contexts" (Buckingham Shum & Ferguson, 2012, p. 4), no studies were identified that examine not only the technical challenges of implementing learning analytics methods but also include an in-depth analysis of the learning contexts in the field of peer assessment. This research aims to fill this gap.

2. RESEARCH DESCRIPTION

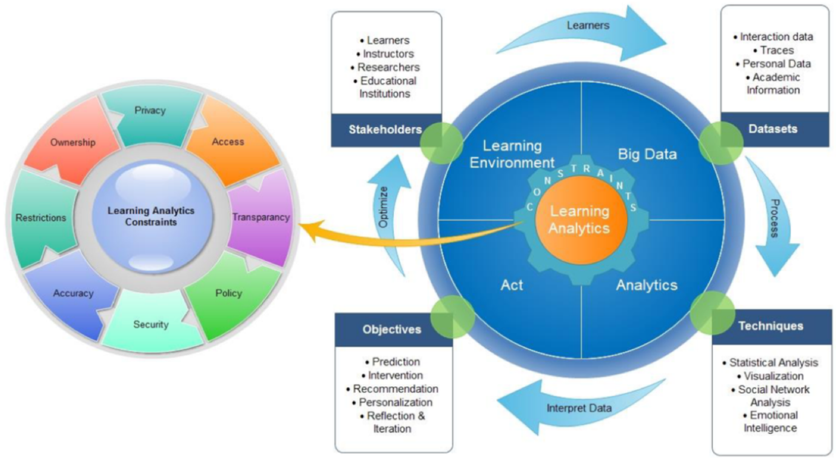
This section introduces the research questions and two of the studies that will be carried out.

2.1. Research Questions and Methodology

Three research questions were developed to guide this research project: 1) How can LA be used to better understand peer assessment learning in higher education to inform teachers and learners? 2) Which LA methods are useful on peer assessment data? How are they useful? 3) How can context data be integrated into the LA analysis?

Two methodological frameworks, which highlight aspects of LA and assessment that need be considered, will guide the development of the research design: *Learning Analytics - Principles and Constraints*

Framework (LA-PCF) (Khalil & Ebner, 2015; see Figure 1) and *Assessment Analytics Framework* (AAF)



(Papamitsiou & Economides, 2016; see Figure 2).

LA-PCF was developed based on previous frameworks proposed by Clow (2012), Chatti et al. (2012), and Greller & Drachsler (2012), and it comprises the LA Life Cycle that describes "proceeding steps, starting from the learning environment and ending with the appropriate intervention" and the LA Constraints that represents limitations of LA research (Khalil & Ebner, 2015, p.1327). The LA Life Cycle model is divided into four main stages: 1) Learning Environment that focuses on how data produced in learning environments can be used to benefit stakeholders; 2) Big Data that indicates the different types of data; 3) Analytics that describes various LA techniques which can be applied to analyse the data; 4) Act which has interpretation of analytics results in focus and use them to optimize LA objectives (Khalil & Ebner, 2015). The LA Constraints model represents aspects of LA implementation that should be taken into consideration, such as Privacy, Transparency, and Ownership (Khalil & Ebner, 2015). Even though the

Figure 1: Learning Analytics - Principles and Constraints Framework (LA-PCF) (from Khalil & Ebner, 2015, p. 1333)

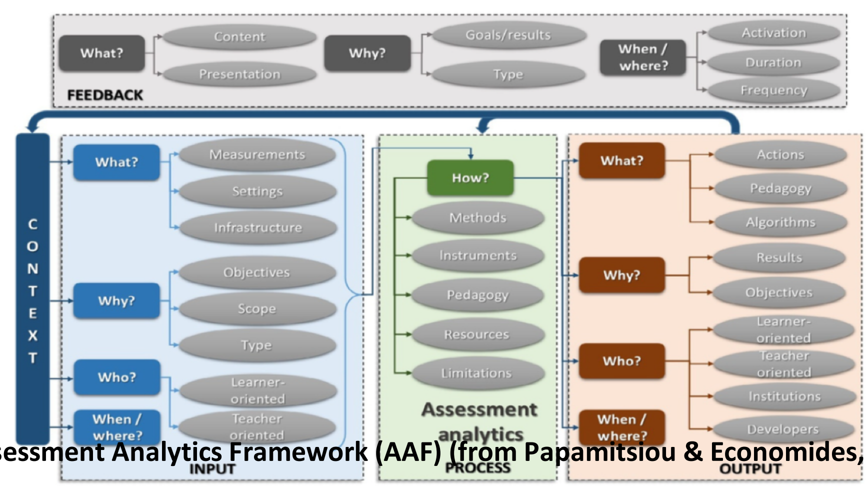


Figure 2. Assessment Analytics Framework (AAF) (from Papamitsiou & Economides, 2016, p. 125)

consequences of the use of LA methods, especially in the areas of student retention and assessment, can significantly influence higher education institutions, ethical and privacy issues are rarely reflected on in LA research (Drachsler et al., 2015).

AAF (Papamitsiou & Economides, 2016) was developed to consolidate learning and assessment research with learning analytics, and consists of four main parts: 1) input that describes the input parameters used for the analytics; 2) process that includes the ways in which data is analyzed and interpreted; 3) output that shows what is the analytics outcome and for which stakeholder should it be available; and 4) feedback that describes not only the feedback given to the final user, but also the feedback iterative process. The concept of context comes from the learning sciences and is an important part of the AAF (Papamitsiou & Economides, 2016). It provides additional information about "the situation of an entity (eg, person, place or object)", which is especially important in the real-life application of the assessment analytics (Papamitsiou & Economides, 2016, p. 125).

2.2. Peergrade Tool

In this research a peergrading tool developed at the Technical University of Denmark, Peergrade, will be used in various courses to implement peer assessment.

Peergrade is "a free online platform to facilitate peer feedback sessions with students"[1] (Peergrade, n.d.). It can be integrated with the most popular LMS platforms such as Moodle and Canvas. The platform not only enables giving feedback to the other students, but also creates a feedback loop, in which the feedback is evaluated by the person who was assessed. The agreement between the graders is calculated, so that the teacher can intervene in the case of a high discrepancies between the grades. There is also data on how much time students spent on giving feedback. Many kinds of files can be uploaded for grading including PDFs, videos, etc. The assignments can be weighted and it is possible to give feedback anonymously.

2.3. Study 1: Big Data Analysis

Peergrade is a partner with Centre for the Science of Learning & Technology (SLATE) at the University of Bergen. The anonymised data from the Peergrade tool has been made available for a retrospective analysis in the SLATE's big data infrastructure. The up to date data set consists of data from hundreds educational institutions that use Peergrade.

The analysis of this big data set focuses on the usefulness of various LA methods on the peer assessment data for helping with the understanding of learning processes. The focus of this study is both learning-centric analytics as well as learner-centric analytics, so not just student behaviour, but also the content and produced learning artefacts will be examined (Stein, 2012). Different LA methods are used to analyse different aspects of learning, such as a predictive analysis of the student performance based on the peer assessment data, and a natural language processing analysis of the content of the students' feedback.

[1] Free for the Basic and a pricing model for additional features.

Since there is little context data available for the analysis, the results from this analysis will be exploratory in nature and will inform the research in Study 2.

2.4. Study 2: Context Data Integration

Study 2 focuses on peer assessment in higher education, especially the use of LA in order to understand how the context data from the setting in which the tool is used, together with the Peergrade data, can provide insights into the role of peer assessment in learning and teaching[2]. In focus are aspects such as learning design (e.g., Peergrade assignments), the way students are trained in giving feedback, how peer assessment is introduced, the quality of the peer feedback and how it changes over time. The goal is to look beyond just the Peergrade data (informed by what is learned in Study 1), and to include contextual data such as focus interviews, videos, etc., as well as to analyze the perception and usefulness of the visualizations presented to the teacher and the learner.

Kristiania University College (HK) collaborates with SLATE in a project using Peergrade. As the collaboration is planned for the next couple of years, there is a possibility to study changes in the learning design and thus perform an iterative process, where based on the LA results interventions are undertaken, analytics are improved, and the cycle is repeated on a new group of students (Clow, 2012). In the fall 2017 semester four HK courses in a variety of discipline are using the Peergrade tool and data generated will be used in the Study 2.

2.5. Current Status of the Research

Two extensive literature reviews on LA have been previously carried out (Misiejuk, 2017; Misiejuk & Wasson, forthcoming). These will be extended with a in-depth revision focused on learning design, teacher inquiry, and peer assessment.

The data is available for the Study 1 in SLATE's big data infrastructure. The exploratory analysis on the big data set has begun with cleaning the data and exploring the database structure. Preliminary natural language processing analysis and simple descriptive analytics have been carried out to map out the kinds of questions that can be answered with this dataset (e.g., sentiment analysis, temporal analysis). The next steps are to crystalize the research questions and methods to address them.

The first courses using Peergrade at the HK are ongoing during fall 2017 semester. Data sets for these courses have been made available. Since Study 2 is part of a larger project, the design of my study will be further developed, and in particular the relationship of LA, learning design, and teacher inquiry.

---

[2] This study is currently being designed and will most likely be part of a larger protect that will be looking at the relationship between learning analytics, learning design, and teacher inquiry (Mor et al., 2015; Wasson et al., 2016). For this reason the study description is general.

REFERENCES

Ashenafi, M.M., Ronchetti, M., & Riccardi, G. (2016). Predicting Student Progress from Peer-Assessment Data. In *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 270-275).

Buckingham Shum, S. & Ferguson, R. (2012). Social Learning Analytics. *Educational Technology & Society*, 15(3), 3-26.

Chatti, M.A., Dyckhoff, A.L., Schroeder, U., & Thüs, H. (2012). A Reference Model for Learning Analytics. *International Journal of Technology Enhanced Learning*, 4(5-6), 318-331.

Clow, D. (2012). The Learning Analytics Cycle: Closing the Loop Effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134-138).

de Alfaro, L. & Shavlovsky, M. (2016). Dynamics of Peer Grading: An Empirical Study. In *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 62-29).

Drachsler, H., Hoel, T., Scheffel, M., Kismihok, G., Berg, A., Ferguson, R., . . . Manderveld, J. (2015). Ethical and Privacy Issues in the Application of Learning Analytics. In *Proceedings of the 5th International Conference on Learning Analytics and Knowledge* (pp. 390-391).

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., . . . Vuorikari, R. (2016). *Research Evidence on the Use of Learning Analytics - Implications for Education Policy*. R. Vuorikari, J. Castano Munoz (Eds.). Joint Research Centre Science for Policy Report.

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Journal of Educational Technology & Society*, 15(3), 42.

Jiang, S., Williams, A.E., Schenke, K., Warschauer, M., & O'Dowd, D.K. (2014). Predicting MOOC Performance with Week 1 Behavior. In *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 273-275).

Khalil, M., & Ebner, M. (2015). Learning Analytics: Principles and Constraints. In *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 1326-1336).

Kulkarni, C.E., Bernstein, M.S., & Klemmer, S.R. (2015). PeerStudio: Rapid Peer Feedback Emphasizes Revision and Improves Performance. In *Proceedings of the 2nd ACM Conference on Learning at Scale* (pp. 75-84).

Kulkarni, C.E., Socher, R., Bernstein, M.S., & Klemmer, S.R. (2014). Scaling Short-Answer Grading by Combining Peer Assessment with Algorithmic Scoring. In *Proceedings of the 1st ACM Conference on Learning at Scale* (pp. 99-108).

Li, H., Xiong, Y., Zang, X., L. Kornhaber, M., Lyu, Y., Chung, K.S., & K. Suen, H. (2016). Peer Assessment in the Digital Age: A Meta-Analysis Comparing Peer and Teacher Ratings. *Assessment & Evaluation in Higher Education*, 41(2), 245-264.

Misiejuk, K. (2017). *Mapping the Field of Educational Data Sciences. Analysis of the Proceedings from Educational Data Mining, Learning Analytics and Knowledge, and Learning at Scale* (Unpublished master's thesis). Humboldt University of Berlin, Germany.

Misiejuk K. & Wasson B. (forthcoming 2018). *State-of-the-Field on Learning Analytics and Knowledge*. Bergen, Norway: The Centre for the Science of Learning & Technology (SLATE).

Mor, Y., Ferguson, R., & Wasson, B. (2015). Learning design, teacher inquiry into student learning and learning analytics: A call for action. *British Journal of Educational Technology*, 46(2), 221-229.

Papamitsiou, Z., & Economides, A. (2016). An Assessment Analytics Framework (AAF) for Enhancing Students' Progress. In S. Caballe & R. Clariso (Eds.), *Formative Assessment, Learning Data Analytics and Gamification: In ICT Education* (pp. 117-133). Morgan Kaufmann.

Patchan, M.M., Schunn, C.D., & Clark, R.J. (2017). Accountability in Peer Assessment: Examining the Effects of Reviewing Grades on Peer Ratings and Peer Feedback. *Studies in Higher Education*, 1-16.

Peergrade, (n.d.) In *Peergrade: Official Website.* Retrieved November 6, 2017, from: https://www.peergrade.io.

Piech, C., Huang, J., Chen, Z., Do, C., Ng, A., & Koller, D. (2013). Tuned Models of Peer Assessment in MOOCs. In *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 153-160).

Raes, A., Vanderhoven, E., & Schellens, T. (2015). Increasing Anonymity in Peer Assessment by Using Classroom Response Technology within Face-to-Face Higher Education. *Studies in Higher Education*, 40(1), 178-193.

Raman, K., & Joachims, T. (2015). Bayesian Ordinal Peer Grading. In *Proceedings of the 2nd ACM Conference on Learning at Scale* (pp. 149-156).

Sajjadi, M.S., Alamgir, M., & von Luxburg, U. (2016). Peer Grading in a Course on Algorithms and Data Structures: Machine Learning Algorithms do not Improve over Simple Baselines. In *Proceedings of the 3rd ACM Conference on Learning at Scale* (pp. 369-378).

Stein, Z. (2012, August 13). *Learning Analytics and the Learning Sciences*. Educause Webinar. http://www.educause.edu/node/267647.

Tritz, J., Michelotti, N., Shultz, G., McKay, T., & Mohapatra, B. (2014). Peer Evaluation of Student Generated Content. In *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (pp. 277-278).

Xing, W., Wadholm, B., & Goggins, S. (2014). Learning Analytics in CSCL with a Focus on Assessment: an Exploratory Study of Activity Theory-Informed Cluster Analysis. In *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (pp. 59-67).

Xiong, W. & Litman, D. (2013) Evaluating Topic-Word Review Analysis for Understanding Student Peer Review Performance. In *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 200-207).

Xiong, W., Litman, D., & Schunn, C.D. (2010). Assessing Reviewers Performance Based on Mining Problem Localization in Peer-Review Data. In *Proceedings of the 3rd International Conference on Educational Data Mining* (pp. 211-220).

Vogelsang, T., & Ruppertz, L. (2015). On the Validity of Peer Grading and a Cloud Teaching Assistant System. In *Proceedings of the 5th International Conference on Learning Analytics and Knowledge* (pp. 41-50).

Vozniuk, A., Holzer, A., & Gillet, D. (2014). Peer Assessment Based on Ratings in a Social Media Course. In *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (pp. 133-137).

Wahid, U., Chatti, M.A., Schroeder, U. (2016). The State of Peer Assessment: Dimensions and Future Challenges. *International Journal on Advances in Systems and Measurements*, 9(3-4), 242-252.

Wasson, B., Hanson, C., & Mor, Y. (2016). Grand Challenge Problem 11: Empowering Teachers with Student Data. In *Grand Challenge Problems in Technology-Enhanced Learning II: MOOCs and Beyond* (pp. 55-58). Springer International Publishing.

Towards Teacher-AI Hybrid Systems

Kenneth Holstein
Human-Computer Interaction Institute, Carnegie Mellon University
kjholste@cs.cmu.edu

ABSTRACT: AI-powered educational software, such as intelligent tutoring systems (ITSs), is commonly designed to enhance student learning. However, such software is not typically designed to effectively collaborate with human teachers. The present work explores how AI and teachers might leverage and amplify one another's complementary strengths to achieve outcomes greater than either could achieve alone. Key contributions of this research, so far, include (1) an initial, broad exploration of K-12 teachers' needs and desires for real-time support in ITS classrooms; (2) the first exploration of the affordances of smart glasses to support orchestration of personalized classrooms, resulting in a prototype called Lumilo; (3) a new prototyping method for real-time teacher support tools; and (4) a classroom experiment evaluating the effects of teacher/AI co-orchestration, supported by Lumilo, on teacher and student behavior and student learning outcomes in classrooms using ITSs.

Keywords: teaching analytics, classroom orchestration, cognitive augmentation, human-machine intelligence, human-in-the-loop, teachers, explainability, design methods, co-design

1 INTRODUCTION

To facilitate more personalized learning, AI-powered educational software is increasingly being used in K-12 classrooms (Pane, Steiner, Baird, Hamilton, & Pane, 2017). Yet teachers struggle with what their new roles in such classrooms should be (Holstein, McLaren, & Aleven, 2017b). One class of AI-powered educational software, intelligent tutoring systems (ITSs), have been shown through several meta-analyses to significantly enhance student learning when used in classrooms, compared with other forms of educational technology and traditional classroom instruction (e.g., Kulik & Fletcher, 2016). These systems provide step-by-step feedback and guidance to students – tailoring instruction to individual learners as they work through problem-solving activities at their own pace. In turn, ITSs also free up the teacher to provide more one-on-one support to students who may benefit from it the most. Yet orchestrating personalized learning also poses unique challenges for teachers, who are tasked with monitoring classes working on a range of educational activities, and prioritizing help across students given limited time (Holstein, McLaren, & Aleven, 2017b).

Over a decade ago, Yacef proposed a reframing of intelligent tutoring systems as "intelligent teaching assistants" (ITAs): systems designed with the joint objectives of helping human teachers teach and helping students learn (rather than only the latter of these objectives, as is typical of ITSs) (Yacef, 2002). Other researchers have since proposed similar directions, with a focus on optimizing student learning by leveraging complementary strengths of human and AI instruction (e.g., Ritter, Fancsali, Berman, & Yudelson, 2014). That is, ITSs might be more effective if they could adaptively enlist the help of human teachers (c.f. Kamar, 2016), in situations teachers may be better suited to handle. While there has been some work on real-time teacher support tools for ITS classrooms since the vision of ITAs was introduced, little work has explored teachers' actual *needs* and *desires* for such support, or how human instruction might most effectively be combined with AI instruction.

In my current and proposed work, I investigate how teachers might best be supported by AI, how AI might best be supported by teachers, and how AI and human teachers might leverage one another's complementary strengths to achieve outcomes greater than either could achieve alone.

2 PRIOR WORK

2.1 Exploring K-12 teachers needs and desires for real-time analytics

Whereas real-time support tools for teachers, such as dashboards, have become popular with many learning technologies (e.g., Martinez-Maldonado, Clayphan, Yacef, & Kay, 2015; Mavrikis, Gutierrez-Santos, & Poulovassilis, 2016), we are not aware of projects (in academic research or in industry) that have conducted a broad investigation of teachers' actual *needs* for real-time support (i.e., one that is not tied to an existing prototype or current technical feasibility, such as the current availability of data or measurement techniques (Holstein, McLaren, et al., 2017b; Rodriguez-Triana et al., 2017)). Furthermore, work on real-time support tools for personalized classrooms, more broadly, has tended to focus on designing tools for *university-level* instructors.

To better understand K-12 teachers' information needs for real-time analytics, my colleagues and I conducted a series of interviews and design studies with ten middle school math teachers (across 5 schools and 5 school districts in Pittsburgh and surrounding areas). For example, in a generative card sorting exercise, we asked teachers what "superpowers" they would want during ITS class sessions, to help them do their jobs. Overall, this exercise revealed that the analytics commonly generated by existing teacher dashboards and reporting systems for ITSs rarely align with those that teachers expect to be most useful (Holstein, McLaren, et al., 2017b). During this card sorting exercise, teachers also generated the idea of being able to see information about individual students "floating over their heads", directly within the physical classroom space. In a follow-up series of concept generation and validation studies, we found that teachers preferred *wearable* awareness tool designs (c.f., Quintana, Quintana, Madeira, & Slotta, 2016) that allowed them to keep their heads up, and their eyes focused on the classroom. Teachers emphasized that handheld real-time dashboards may compete for attention with some of the most useful real-time information in the classroom: student body language and other cues that would not be captured by a dashboard alone (Holstein, McLaren, et al., 2017b). In particular, teachers gravitated towards the idea of wearing eyeglasses that could grant them a private view of actionable information about their students, embedded through the classroom space (c.f., Alavi, Dillenbourg, & Kaplan, 2009), without revealing sensitive data to students or their peers (c.f. Jivet, Scheffel, Drachsler, & Specht, 2017).

2.2 Opening up an ITS development environment for extensible student modeling

To support the development of learning analytics tools for use with ITSs (e.g., teacher- and student-facing dashboards), my colleagues and I are substantially extending the existing CTAT and TutorShop architecture for ITS authoring and deployment (Holstein, Yu, et al., 2018c). The extended technical architecture, *CTAT/TutorShop Analytics* (CT+A), supports the authoring, sharing, and re-use of a broad and open range of student modeling techniques and analytics, for use in running ITSs (i.e., to drive adaptive tutoring behavior) and/or external learning analytics tools.

2.3 Co-designing wearable cognitive augmentation for K-12 teachers

2.3.1 Lumilo: a real-time awareness tool for personalized K-12 classrooms

Building on findings from our initial user-centered design research with K-12 teachers, briefly summarized in (2.1), my colleagues and I next conducted a series of iterative, participatory design studies with a total of 16 middle school math teachers (across 9 schools and 6 school districts in Pittsburgh and surrounding areas). We began with storyboarding, lo-fi prototyping, and participatory sketching sessions, to validate teachers' desires for real-time analytics, further probe underlying needs, and explore how teachers envisioned actually *using* this information during a class session (Holstein, Hong, Tegene, McLaren, & Aleven, 2018a). We also further explored the idea of "teacher smart glasses" further, to understand their unique affordances for monitoring personalized classes. After a Wizard-of-Oz'd mid-fidelity prototyping phase, we created a fully-functional prototype of a mixed-reality smart glasses based orchestration tool called Lumilo, capable of interfacing with a broad range of ITSs (Holstein, Hong, et al., 2018a). Using the CT+A architecture (2.2), we developed an initial set of automated detectors, using established student modeling techniques (Desmarais & Baker, 2012). To facilitate iterative prototyping of both real-time analytics and their visualizations, we also developed a novel prototyping method (discussed in the next section).

2.3.2 Replay Enactments: a prototyping method for real-time teacher support tools

To prototype the experience of using Lumilo in a classroom, we developed a new prototyping method for real-time teacher support tools: Replay Enactments (Holstein, Hong, et al., 2018a). Like other recently proposed methods in Learning Analytics (Martinez-Maldonado et al., 2016; Mavrikis et al., 2016), Replay Enactments (REs) involve replaying log data from student-software interactions, to prototype real-time analytics and their visualizations. REs go beyond these approaches by emphasizing embodied role playing and simulation exercises in physical classroom spaces, in the spirit of recent HCI methods, such as User Enactments, for prototyping radically new experiences (Odom et al., 2012). In contrast to User Enactments, however, REs prototype experiences using authentic data and algorithms, unfolding over time. Doing so facilitates observation of the interplay between teacher and machine judgments, including the UX impact of a prototype's false positives and false negatives (Dove, Halskov, Forlizzi, & Zimmerman, 2017).

2.4 Investigating relationships between teacher attention, student behavior, and student learning

To facilitate the discovery of relationships between out-of-software interactions (e.g., teacher-student interactions) and student learning within educational software in blended learning environments, we developed a new log replay tool: the Spatial Classroom Log Explorer (SPACLE) (Holstein, McLaren, & Aleven, 2017c). Using SPACLE, my colleagues and I found that students' *mere awareness* of being monitored by a teacher may contribute to student engagement and learning. We also found early evidence that, in classrooms not using a teacher awareness tool, students who exhibit patterns of "help avoidance" (Aleven, Roll, McLaren, & Koedinger, 2016) *within* educational software also tend to receive less teacher attention (Holstein, McLaren, & Aleven, 2017a).

We enhanced Lumilo to collect moment-by-moment data on a teacher's activity within the physical classroom, including gaze, position, orientation, and physical proximity to various classroom hotspots. We then conducted an in-lab experimental study, using REs, with results suggesting that Lumilo measurably directs teachers' attention towards students who would go on to exhibit lower performance on a posttest, compared with business-as-usual (Holstein, Hong, et al., 2017). Follow-up analyses using causal path modeling suggested that this effect was explained largely by Lumilo's alerts about student "unproductive persistence", or "wheel-spinning" (Kai, Almeda, Baker, Shechtman, Heffernan, & Heffernan, 2018), in educational software. To a lesser extent, this effect also appears to arise because Lumilo directs teachers' attention to students who less effectively regulate their own help-seeking behavior within the software, compared with business-as-usual (where help-avoidant students are relatively neglected) (Holstein, McLaren, et al., 2017a).

3 ONGOING WORK

3.1 A classroom experiment to study the effects of real-time teacher analytics

My colleagues and I have recently run in-vivo classroom experiments (with 286 middle school students, across 18 classrooms and 8 teachers) to investigate the effects of providing teachers with real-time analytics about student learning, metacognition, and behavior, on (1) teacher behavior; (2) student behavior and performance; and (3) students' out-of-software learning gains (Holstein, McLaren, & Aleven, 2018b). Among other findings, the results indicate that a teacher's use of *Lumilo* had a positive impact on student learning, compared with both business-as-usual and simpler classroom monitoring support. Real-time teacher analytics served as an equalizing force in the classroom: narrowing the gap in learning outcomes across students of varying prior ability.

Prior work has found that providing teachers with real-time notifications about student performance can direct their attention to *low-performing* students, resulting in local performance improvements (e.g., Martinez-Maldonado et al., 2015). Other recent work has begun systematically investigating how teachers use real-time progress and performance analytics in blended classrooms (e.g., Molenaar & Knoop-van Campen, 2017). However, the present work is the first experimental study showing that real-time teacher analytics can enhance students' *learning outcomes*.

We are currently analyzing data collected from this classroom study to better understand how the real-time analytics presented by Lumilo influenced teacher-student interactions. These classroom experiments also provided an opportunity to gather students' perspectives on the current design of Lumilo. As such, we are also currently analyzing this design feedback, to inform the design of future teacher-AI hybrid tools that can more effectively serve *students'* needs and desires.

4 PROPOSED WORK

4.1 Opening up the black box: Supporting teacher interpretation of AI inferences in teacher-AI hybrid systems.

Early in our design research with teachers, it became clear that teacher *autonomy* is a central issue in the design of teacher-AI hybrid systems (Holstein et al., 2017b; 2018a). While on the one hand teachers have often requested more direct *decision support* than is commonly offered by teacher dashboards (e.g., in the form of real-time action recommendations), especially in the face of limited time, teachers have also revealed strong discomfort with AI systems that they perceive to be "telling them what to do". During the iterative design of Lumilo, we began to explore how teacher-AI hybrid systems might effectively balance teacher autonomy with this desire for real-time decision support.

Prototyping studies with teachers (Holstein et al., 2018a) suggested that teachers' ability to *interpret* inferences and recommendations made by the AI was key not only in facilitating teacher *trust* in the AI, but also in empowering teachers to override the AI's decisions, if need be (c.f. Kamar, 2016). However, "interpretability" is a very broad notion, and in general, little is known about the effects of different forms of AI interpretability on end-users' trust, feelings of autonomy, and decision-making (Doshi-Velez & Kim, 2017; Lipton, 2016). In our prototyping studies, for example, we discovered that teachers were not particularly concerned with the *intelligibility* of an AI model (e.g., visualizations aimed at helping teachers understand *how* the AI arrived at an inference, or how the AI model was learned/trained in the first place). Instead, teachers strongly preferred *post-hoc explanations* of AI inferences (Lipton, 2016), such as curated snippets of a student's interactions with the software that could help "corroborate" a given claim made by the AI (Holstein et al., 2018a).

I propose to build on these initial investigations by systematically, empirically investigating a broader design space of mechanisms for AI to explain their own inferences (c.f. Doshi-Velez & Kim, 2017; Lipton, 2016) in teacher-AI hybrid systems, across multiple relevant evaluation criteria (e.g., effects on teacher trust in the system, as well as teachers' ability to make more informed decisions about when to override or modify AI decisions/recommendations). I expect these investigations will ultimately help pave the way for more effective and desirable partnerships between human teachers and AI systems. For example, greater interpretability may help offset otherwise harmful effects of undesirable algorithmic biases, which commonly arise in data-driven intelligent systems (Doshi-Velez & Kim, 2017; Kamar, 2016; Lipton, 2016). In addition, enabling teachers to better *understand* AI inferences may be a step towards enabling teachers to interactively provide *feedback* to these systems, to improve their usefulness within a specific classroom context (Kamar, 2016).

4.2 "Humble" AI in education: AI tutors that recognize their own limitations

The phenomenon of "unproductive persistence" (Kai et al., 2018), in AI-powered educational software can be understood as the software reaching its own pedagogical limitations (Holstein, McLaren, et al., 2017b; Käser & Gross, 2016). That is, any situation where a *learner* persists in educational software without mastering the material can also be understood as the *software* unproductively persisting in a particular teaching strategy. As such, in scenarios where a teacher is present, unproductive persistence may be viewed as a critical opportunity for the AI to (humbly) "pass control" to the teacher. Towards the design of more “humble” intelligent tutoring systems (c.f. Baker, 2016), I propose to: (1) mine pre-existing datasets, using multiple measures of “unproductive persistence” and models of other constructs (Desmarais & Baker, 2012), to better understand the *causes* of unproductive persistence within ITSs; (2) leverage findings from these analyses to develop *earlier* and *more accurate* computational methods to distinguish “unproductive” from “productive” persistence in ITSs (c.f. Kai et al., 2018); and (3) building on the work in (1), as well as section 4.1, design a system that supports teachers in interpreting the inferences made by these computational methods, and in effectively responding to instances of unproductive persistence.

REFERENCES

Alavi, H. S., Dillenbourg, P., & Kaplan, F. (2009). Distributed Awareness for Class Orchestration. In *EC-TEL* (pp. 211–225).

Aleven, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2016). Help Helps, but only so Much: Research on Help Seeking with Intelligent Tutoring Systems. *IJAIED*, *26*(1), 205–223.

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *IJAIED*, *26*(2), 600-614.

Desmarais, M. C., & Baker, R. S. J. D. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *UMUAI*, *22*, 9–38.

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In *CHI* (pp. 278–288). ACM.

Doshi-Velez, F., & Kim, B. (2017). A roadmap for a rigorous science of interpretability. *arXiv preprint arXiv:1702.08608*.

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018a). Co-designing Wearable Cognitive Augmentation for K-12 Teachers. To appear in *LAK’18*. ACM.

Holstein, K., McLaren, B. M., & Aleven, V. (2017a). Informing the design of teacher awareness tools through Causal Alignment Analysis. Under Review.

Holstein, K., McLaren, B. M., & Aleven, V. (2017b). Intelligent tutors as teachers’ aides: exploring teacher needs for real-time analytics in blended classrooms. In *LAK’17* (pp. 257–266). ACM.

Holstein, K., McLaren, B. M., & Aleven, V. (2017c). SPACLE: investigating learning across virtual and physical spaces using spatial replays. In *LAK’17* (pp. 358–367). ACM.

Holstein, K., McLaren, B. M., & Aleven, V. (2018b). Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. Under Review.

Holstein, K., Yu, Z., Sewall, J., Popescu, O., McLaren, B. M., & Aleven, V. (2018c). Opening up an ITS development environment for extensible student modeling. Under Review.

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: pitfalls of learning analytics dashboards in the educational practice. In *EC-TEL* (pp. 82-96). Springer, Cham.

Kai, S., Almeda, V.A., Baker, R.S., Shechtman, N., Heffernan, C., Heffernan, N. (2018). Modeling wheel-spinning and productive persistence in Skill Builders. To appear in *JEDM*.

Kamar, E. (2016). Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence. In *IJCAI* (pp. 4070-4073).

Käser, T., & Gross, M. (2016). When to stop ? Towards Universal Instructional Policies. In *LAK’16* (pp. 289–298). ACM.

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *RER*, *86*(1), 42–78.

Lipton, Z. C. (2016). The mythos of model interpretability. *arXiv preprint arXiv:1606.03490*.

Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2015). MTFeedback: Providing Notifications to Enhance Teacher Awareness of Small Group Work in the Classroom. *IEEE TLT*, *8*(2), 187–200.

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2016). LATUX: An iterative workflow for designing, validating and deploying learning analytics visualisations. *JLA*, *2*(3), 9–39.
Mavrikis, M., Gutierrez-Santos, S., & Poulovassilis, A. (2016). Design and Evaluation of Teacher Assistance Tools for Exploratory Learning Environments. In *LAK* (pp. 168–172). ACM.
Molenaar, I., & Knoop-van Campen, C. (2017). Teacher dashboards in practice: Usage and impact. In *EC-TEL* (pp. 125-138). Springer, Cham.
Odom, W., Zimmerman, J., Davidoff, S., Forlizzi, J., Dey, A., & Lee, M. K. (2012). A Fieldwork of the Future with User Enactments. In *DIS* (pp. 338–347). ACM.
Pane, J., Steiner, E., Baird, M., Hamilton, L., & Pane, J. (2017). *Informing Progress: Insights on Personalized Learning Implementation and Effects*. RAND Corporation.
Quintana, R., Quintana, C., Madeira, C., & Slotta, J. D. (2016). Keeping Watch: Exploring Wearable Technology Designs for K-12 Teachers. In *CHI* (pp. 2272–2278). ACM.
Ritter, S., Fancsali, S. E., Berman, S., & Yudelson, M. (2014). Towards Integrating Human and Automated Tutoring Systems, In *EDM* (pp. 626–627).
Rodríguez-Triana, M. J., Prieto, L. P., Vozniuk, A., Boroujeni, M. S., Schwendimann, B. A., Holzer, A., & Gillet, D. (2017). Monitoring, awareness and reflection in blended technology enhanced learning: a systematic review. *IJTEL*, *9*(2-3), 126-150.
Yacef, K. (2002). Intelligent Teaching Assistant Systems. In Kinshuk (Ed.), *IEEE* (pp. 136–140).

Understanding the design of Learning Analytics for Students

Liliana Farias Herrera
University of Hong Kong
lilians@connect.hku.hk

ABSTRACT: Current learning analytics (LA) commonly provide to learners descriptive information and feedback regarding their learning behavior and academic performance. However, they generally do not receive formative feedback that guides reflection in making sense of the LA outcomes for improving the learning process strategies on future learning opportunities. Recently, there are diverse projects that focus on finding different strategies to close the feedback loop, but many of these are teacher-centered, with few projects that focused from the learners' perspectives. There is a need to provide students with formative feedback informed by learning analytics that empower their agency and decision making to improve their learning process as self–directed learners. This research project aims to investigate how LA can be designed and implemented effectively to support students to make sense of LA formative feedback on their learning outcomes, learning strategies and process.

Keywords: Feedback, learning design, learning analytics, self-directed learning.

1 INTRODUCTION

Technology Enhanced Learning (TEL) has stimulated the offerings of and demands for blended and online courses by stimulating the increased the availability of students' digital footprint. This augmented possibilities to analyze big data through different lenses (Duval, 2011) in order to understand how to help students to succeed (Bodily & Verbert, 2017; Siemens & Long, 2011). In turn, the growing diversity of LA practitioners has led to greater diversities in the perspectives underpinning the learning analytics tools, reporting systems, visualizations and dashboards available. Some describe LA as a statistical model or methodological application that identifies, diagnose and predict patterns based on data sets collected from digital learning environments or assessments (Lockyer, Heathcote, & Dawson, 2013; McKenney & Mor, 2015a; Mor, Ferguson, & Wasson, 2015). Others recognize that LA techniques supports the design of educational reports by providing relevant information through visualizations or dashboards (Bodily & Verbert, 2017; Epp & Bull, 2015). LA is also referred to as a collection of different processes (Gasevic, Dawson, Mirriahi, & Long, 2015; Siemens & Long, 2011), or a new research field that is connects data scientists and educators (Knight, Shum, & Littleton, 2014; Siemens, 2012).

Therefore, there is no common agreement among practitioners on how to approach and implement learning analytics that is underpinned by a pedagogically sound approach with an emphasis on the learning (Persico & Pozzi, 2015; Siemens, 2012). The pedagogical contextualization of learning data provides a meaningful information to ease decision making to improve the quality of learning as well as the design of learning experiences (Bakharia et al., 2016; Epp & Bull, 2015).

As the LA field matures, it could play a key role in addressing pedagogical problems (Bakharia et al., 2016) of teachers and learners by allowing the design process of self-directed learning opportunities and reporting relevant information grounded on data of the learning process (Bakharia et al., 2016; Corrin & de Barba, 2014; Sharples et al., 2016; Wise, Vytasek, Hausknecht, & Zhao, 2016). Many of the empirical LA research projects that support teaching and learning are teacher-centred, with few that focus on the learners' perspectives in which students receive meaningful LA formative feedback. There is a substantial and growing interest in exploring student facing-systems (Bodily & Verbert, 2017) for providing LA reports, visualizations and dashboards in order to enhance learners awareness of their current achieved learning outcomes and compare them with the expected ones. For example, open learner models (OLMS), recommender systems, learning analytics dashboards and tutoring systems are those student facing-systems commonly explored by scholars and practitioners (Bodily & Verbert, 2017; Epp & Bull, 2015).

The LA field has made emphasis on the fundamental issues of data collection, management, processing and visualizations without focusing on the learning context (Bodily & Verbert, 2017; Sharples et al., 2016; Wise, 2014). There is a need to have a conceptual framework that provides a common understanding and interpretation of the analytics that can be accessible for use by teachers and learners (Bakharia et al., 2016; Lockyer et al., 2013). Currently, available LA reports are generally teacher-centred (Law, Li, Farias Herrera, Chan, & Pong, in press), there is a need to investigate how student-led learning analytics empower their agency and decision making to improve their learning process as self-directed learners.

The present study aim to addresses a relatively unexplored area in the evolving field of learning analytics for learners. The rest of this research proposal is organized as follow. Section 2 provides a brief background of the research problem by stating the need of LA for learners. Section 3 describes a brief literature review. Section 4 explains the research goals and questions to be address, and section 5 describes the preliminary sketch of the methodology.

2 STATEMENT OF THE RESEARCH PROBLEM

Learning analytics for students commonly provide descriptive information and feedback of their academic performance and behavior within digital learning environments, in which the LA reports, dashboards or visualizations are representations from teachers' perspective (Corrin & de Barba, 2014; Epp & Bull, 2015; Kitto, Lupton, Davis, & Waters, 2016; Wise, 2014; Wise et al., 2016). As a result, students do not receive formative feedback that guides reflection in making sense of the LA outcomes for improving the learning process strategies on future learning opportunities (Corrin & de Barba, 2014; Wise, 2014; Wise et al., 2016). Formative feedback informed by LA for students should connect students with teacher's instructional intentions through LA dashboards and visualizations. Before the development of such a dashboard for students, we need to consider how students interpret the LA results presented to them. Besides looking only at the statistics, learners probably require additional pedagogical information that eases the process of sense making of the information reflected in LA dashboards or visualizations. There is a need to provide students with formative feedback informed by learning analytics that empower their agency and decision making to improve their learning process as self–directed learners.

3 RELATED WORK

3.1 Integrating LA into teaching and learning practices

The disruptive impact of TEL on teaching and learning has provoked discussions and investigations from different fields with the aim to help people to learn within different learning environments (e.g. formal or informal, online or blended) (Mor et al., 2015; Persico & Pozzi, 2015). For example, fields such as learning design and learning analytics aim to support pedagogical decision making for a meaningful design of learning experiences and achieve expected learning outcomes (Bakharia et al., 2016; Persico & Pozzi, 2015).

Learning design (LD) emerged with the intention to provide guidance and support for those individuals who are committed to enhance teaching and learning through shareable TEL experiences (Law et al., in press; Lockyer, Lori & Dawson, Shane, 2011). As a result, learning design foster the transformation of teachers to become designers of learning experiences by sharing and adoption of best practices (Law et al., in press). Learning design offers LA a pedagogical perspective and vocabulary to contextualize and interpret the data collected (Bakharia et al., 2016). Similarly, LD can be used as a framework to support the design of LA (Lockyer et al., 2013) by enabling teachers to select proper analytics for different pedagogical designs.

During the last years, there is an increase recognition of the benefits in connecting these fields(Bakharia et al., 2016; Lockyer et al., 2013; Lockyer, Lori & Dawson, Shane, 2011; Mor et al., 2015). Therefore, new empirical projects have emerged and are relatively new proposals which most of them have been piloted or still at early stage of research. For example, P.S & Scupelli, P. (2017.) propose a framework in which LA and LD are a powerful synergy to connect stakeholders from different fields to ease their collaborative work and communication. Other projects have focused on designing visualizations, for example Mavriks and Karlalas (2016) designed a learning analytics dashboard for learning designers of e-books with the intention to provide further information that guide them during the pedagogical re-design of an e-book.

In the same way, recent research investigations have focused on fostering teacher decision making through LA and inquiry. It is possible to answer questions of teaching and learning through inquiry-based learning design supported by learning analytics tools (Mor et al., 2015; Persico & Pozzi, 2015). For instance, McKenney and Mor (2015) propose a set of design guidelines to be considered when designing a tool that combines LA, learning design and teacher inquiry, these guidelines propose that a tool should be able to collect data during classroom interactions and provide explicit guidelines on how to interpret and design interventions to improve teaching and learning.

Recently, diverse projects are focused on finding different strategies to close the feedback loop, but many of those are teacher-centered, connecting LD and LA in which teacher inquiry play a mediating role among fields (Karkalas & Mavrikis, 2016; McKenney & Mor, 2015a; Mor et al., 2015a; Persico & Pozzi, 2015). On the other hand, there is a need to identify how to enable learners' inquiry as a strategy to connect learning design and learning analytics from students' perspective.

3.2 Learning analytics supporting students' learning

There is a common concern of preparing individuals ready for the uncertainties of the 21st century. For instance, there is a general view that students need to be able to make effective use of technology and their knowledge to collaborate with others, to solve real world problems, to set their own learning goals, to create and to continue to become self-directed learners (Kozma, 2009; Voogt & Roblin, 2010). These skills are often referred as 21st century skills or competencies (Voogt & Roblin, 2010). As a result, educational institutions are adjusting their teaching and learning strategies with the intention to engage students in developing their 21st century skills to become a lifelong learner (Buckingham Shum & Crick, 2016). There is a diversity of frameworks with the aim to provide guidance and strategies for the implementation of 21st century skills or competencies. Some of the most common frameworks are: (1) Partnership for 21st century skills (http://www.p21.org/); (2) Assessment and Teaching of 21st Century Skills (https://www.atcs.com/) and (3) enGauge (http://pict.sdsu.edu/engauge21st.pdf). These share similar goals such as develop high order thinking skills, communication and collaboration, creativity, critical thinking, citizenship, problem solving in order to prepare students for the knowledge society and to engage them in become self-directed learners (Voogt & Roblin, 2010).

The LA community has raised the concern of developing new solutions to support the implementation of 21st century skills through learning analytics (Buckingham Shum & Crick, 2016; Lockyer et al., 2013). For example supporting students' digital writing skills (Kitto et al., 2016), experiential and collaborative learning (Koh, Shibani, Tan, & Hong, 2016), students' metacognitive behaviour (Kitto et al., 2016; Pardo, Han, & Ellis, 2017); building and sharing knowledge within social networks such as discussion forums (Reich, Tingley, Leder-Luis, Roberts, & Stewart, 2014). On the other hand, many of these research investigations of LA for 21st century skills are focused on supporting some specific stakeholders such as researchers, teachers or administrators. Students also need to be able to set their own learning goals, to collaborate and communicate with others, and be self-directed learners (Buckingham Shum & Crick, 2016).

Learners should be aware of the received feedback for making sense of which further actions are required to be implemented in order to achieve expected outcomes (Corrin & de Barba, 2014; Gasevic, Jovanovic, Pardo, & Dawson, 2017; Kitto et al., 2016; Sadler, 2010; Wise et al., 2016). Recently, there is a substantial and growing body of research in exploring open learner models with the aim to stimulate students' metacognitive and self-directed learning skills by providing customized feedback and visualizations (Bodily & Verbert, 2017; Bull, 2012; Epp & Bull, 2015). On the other hand, it has not been reported how this scaffold students learning process into complex learning outcomes. During the last years, there is a current trend of developing feedback informed by LA dashboards and visualizations for students (Bodily & Verbert, 2017; Epp & Bull, 2015; Sharples et al., 2016), with the aim to increase their awareness and aid their monitoring of their particular learning process. For example, higher education institutions (e.g. UK Open University), scholars (e.g. Corrin & de Barba, 2014) or commercial management systems (e.g Blackboard) are enabling visualizations dashboards and reports to students regarding their academic performance and behavioural data obtained from digital learning environments (Sharples, M., et al., 2016). These LA dashboards and visualizations are designed from the teachers' perspective which sometimes learners identify complex for making sense and design further actions (Bodily & Verbert, 2017).

Other projects have analysed LA tools for supporting self-regulated learning, for example Pardo and colleagues (Gasevic et al., 2017; Pardo et al., 2017; Pardo & Mirriahi, 2017) identified that LA visualizations and dashboards tools support learners to self-regulate their learning process by receiving real-time information of their learning performance while also easing learning designers' pedagogical decisions on how to improve learning experiences based on learners' needs (McKenney & Mor, 2015; Persico & Pozzi, 2015). In the same way, Wise and colleagues show that students can use learning analytics as a stimulus of their self-regulatory cycle (Wise et al., 2016). Hence, learning designers should also be able to identify the pedagogical strategies to integrate LA tools for learners. In summary, recent efforts have been implemented to stimulate students' academic performance through LA for learners, but there is a need to emphasize on learning analytics for students in order to maximize their learning experiences while fostering their self-directed learning and reflection of their performance in order to guide their sense making of future actions.

4 PURPOSE OF THE STUDY

4.1 Research goal and questions

The goal of this research project is mainly focused on LA for learners as key users. This is a novel research on how to provide relevant feedback to enhance students' learning self-directed learning experiences. Therefore, the research goal is to investigate and implement from the learners' perspective how LA supports students to make sense of the formative feedback that LA provides in terms of the learning outcomes, learning strategies and processes. This study looks at the following research questions:

1. Should learning analytics be included as learning design component to provide effective formative feedback to support self-directed learners?
   1.1. What kind of learning analytics would be helpful to the learners for self-monitoring, self-managing or self-modifying their learning process?

2. Would providing the learning design conceptualizations of the course to students be helpful for their ability to make sense of the LA outcomes in order to design further actions in their learning process?
   2.1. How will students interpret the LA results presented to them?
   2.2. Would providing the LD underpinning of the course activities and assessment be helpful to them in optimizing their agency in taking control of their own learning?

3. What kind of LA visualizations and dashboards would be helpful for supporting self-directed learners?

5 METHOD

This project research is based on a case study by using results from quantitative data analysis and qualitative data analysis informed by interviews. The combination of methods provides a better understanding of how students perceive the course learning design by looking at students' attitudes towards learning and compared with their learning performance.

The context of the study is expected to be conducted in undergraduate common core courses that are offering blended learning experiences with self-directed learning opportunities. At the same time, those courses require already to provide some digital visualizations and dashboards to inform students' learning progress or identify teachers who are eager to collaborate during the design of LA visualizations and dashboards.

Within the participant courses, a selection of learning design patterns will be chosen in order to design self-directed learning opportunities for enhancing particular complex learning outcomes. Similarly, based on Based on an assessment of SDL of Costa & Kallick (2004) additional learning supports for SDL will be developed such as rubrics, templates, self-assessments and learning strategies recommendations. It is expected that these opportunities are going to be captured by the digital learning system. From the information collected is expected to design LA visualizations and dashboards that inform learners of their self-directed learning progress. Since this work is an early stage the design of the instruments and data collection still on the analysis phase, for example, a scale requires to be developed in response to a need for a valid and reliable instrument to measure self-directed learning readiness. Therefore, it is expected to use an empirical basis the Self-Directed Learning Readiness Scale (SDLRS) developed by Guglielmino (1997). Based on an assessment of SDL of Costa & Kallick (2004) the interview will be designed. This research is a work in process which will require further adjustments.

REFERENCES

Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gašević, D., Mulder, R., Lockyer, L. (2016). A conceptual framework linking learning design with learning analytics (pp. 329–338). ACM Press. https://doi.org/10.1145/2883851.2883944

Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, *10*(4), 405–418.

Buckingham Shum, S., & Crick, R. D. (2016). Learning Analytics for 21st Century Competencies. *Journal of Learning Analytics*, *3*(2), 6–21.

Bull, S. (2012). Preferred features of open learner models for university students. In *International Conference on Intelligent Tutoring Systems* (pp. 411–421). Springer.

Corrin, L., & de Barba, P. (2014). Exploring students' interpretation of feedback delivered through learning analytics dashboards. In *Proceedings of the ascilite 2014 conference* (pp. 629–633).

Costa, A. L., & Kallick, B. (2004). *Assessment Strategies for Self-Directed Learning*. Corwin Press.

Duval, E. (2011). Attention please!: learning analytics for visualization and recommendation. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 9–17). ACM.

Epp, C. D., & Bull, S. (2015). Uncertainty representation in visualizations of learning analytics for learners: current approaches and opportunities. *IEEE Transactions on Learning Technologies*, *8*(3), 242–260.

Gasevic, D., Dawson, S., Mirriahi, N., & Long, P. D. (2015). Learning analytics–a growing field and community engagement. *Journal of Learning Analytics*, *2*(1), 1–6.

Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics*, *4*(2), 113–128.

Guglielmino, L. M. (1997). Reliability and validity of the self-directed learning readiness scale and the learning preference assessment. *HB Long & Associates, Expanding Horizions in Self-Directed Learning*, 209–222.

Inventado, P. S., & Scupelli, P. (2017). An Online Learning Collaboratory to Address Multidisciplinary Learning Challenges at Scale. *Interaction Design and Architecture (S) Journal-IxD&A*, *33*(EPFL-ARTICLE-231721).

Karkalas, S., & Mavrikis, M. (2016). Towards analytics for educational interactive e-books: the case of the reflective designer analytics platform (RDAP). In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 143–147). ACM.

Kitto, K., Lupton, M., Davis, K., & Waters, Z. (2016). Incorporating student-facing learning analytics into pedagogical practice.

Knight, S., Shum, S. B., & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics*, *1*(2), 23–47.

Koh, E., Shibani, A., Tan, J. P.-L., & Hong, H. (2016). A pedagogical framework for learning analytics in collaborative inquiry tasks: an example from a teamwork competency awareness program (pp. 74–83). ACM Press. https://doi.org/10.1145/2883851.2883914

Kozma, R. (2009). Transforming education: Assessing and teaching 21st century skills. *The Transition to Computer-Based Assessment*, *13*.

Law, N., Li, L., Farias Herrera, L., Chan, A., & Pong, T.-C. (in press). A Pattern Language Based Learni ng Des ign St udio-r esubmi ssion 201 831.pdf. IxD&A Journal.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, *57*(10), 1439–1459.

Lockyer, Lori, & Dawson, Shane. (2011). *Learning Designs and Learning Analytics*. Retrieved from http://dl.acm.org/citation.cfm?id=2090116

McKenney, S., & Mor, Y. (2015). Supporting teachers in data-informed educational design. *British Journal of Educational Technology*, *46*(2), 265–279. https://doi.org/10.1111/bjet.12262

Mor, Y., Ferguson, R., & Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action: Learning design, TISL and learning analytics. *British Journal of Educational Technology*, *46*(2), 221–229. https://doi.org/10.1111/bjet.12273

Pardo, A., Han, F., & Ellis, R. A. (2017). Combining University Student Self-Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance. *IEEE Transactions on Learning Technologies*, *10*(1), 82–92. https://doi.org/10.1109/TLT.2016.2639508

Pardo, A., & Mirriahi, N. (2017). Design, Deployment and Evaluation of a Flipped Learning First-Year Engineering Course. In *The Flipped Classroom* (pp. 177–191). Springer.

Persico, D., & Pozzi, F. (2015). Informing learning design with learning analytics to improve teacher inquiry: Informing LD with LA to improve teacher inquiry. *British Journal of Educational Technology*, *46*(2), 230–248. https://doi.org/10.1111/bjet.12207

Reich, J., Tingley, D. H., Leder-Luis, J., Roberts, M. E., & Stewart, B. (2014). Computer-assisted reading and discovery for student generated text in massive open online courses.

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*(5), 535–550.

Sharples, M., de Roock, R., Ferguson, R., Gaved, M., Herodotou, C., Koh, E., Rienties, B. (2016). Innovating pedagogy 2016: Open University innovation report 5.

Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 4–8). ACM.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 30.

Voogt, J., & Roblin, N. P. (2010). 21st CENTURY SKILLS.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics (pp. 203–211). ACM Press. https://doi.org/10.1145/2567574.2567588

Wise, A. F., Vytasek, J. M., Hausknecht, S., & Zhao, Y. (2016). Developing learning analytics design knowledge in the "middle space": The student tuning model and align design framework for learning analytics use. *Online Learning*, *20*(2). Retrieved from http://olj.onlinelearningconsortium.org/index.php/olj/article/view/783

Posters

SHEILA policy framework – supporting higher education to integrate learning analytics

Yi-Shan Tsai
The University of Edinburgh
yi-shan.tsai@ed.ac.uk

Maren Scheffel
Open Universiteit Nederland
Maren.Scheffel@ou.nl

Dragan Gašević
The University of Edinburgh
Monash University
dragan.gasevic@ed.ac.uk

ABSTRACT: This poster will provide an overview of the development process of the SHEILA policy framework, which is a research product aimed to assist with the development of institutional learning analytics policies. The preliminary findings of our early research activities highlight prominent challenges that threaten the adoption, scalability and sustainability of learning analytics among higher education institutions. This poster calls for a systematic approach to addressing these challenges by presenting a research evidence-based policy framework.

Keywords: Learning analytics; policy; higher education

1 INTRODUCTION

The field of learning analytics (LA), with its associated methods of online student data analysis, is able to provide novel and real-time approaches to assessing critical issues such as student progression and retention. While the use of LA has gained much attention and has been/is being adopted by many higher education institutions (HEIs) in Europe and the other parts of the world, the maturity levels of HEIs in terms of being 'student data informed' are only in the early stages. Literature has identified that the adoption of LA in complex educational systems requires a systematic approach to bring about effective changes (Macfadyen, Dawson, Pardo, & Gaševic, 2014). The SHEILA project aims to support HEIs in the development of institutional policies and strategies for LA by building a policy framework that is based on research evidence.

2 METHODS

The SHEILA policy framework will be developed using the Rapid Outcome Mapping Approach (ROMA) (Macfadyen et al., 2014), based on data collected from direct engagement with stakeholders using various research methods, including group concept mapping (with LA experts, n = 30), interviews (with senior managers, n = 64), surveys (institutional survey, n = 46; staff survey in 4 HEIs, and student survey in 5 HEIs), and focus groups (staff focus groups in 4 HEIs, and student focus groups in 4 HEIs).

3 RESULTS

3.1 Essential features of a LA policy

The group concept mapping activity identified six themes among 99 statements about essential features of a LA policy, including (1) privacy & transparency, (2) roles & responsibilities (of all stakeholders), (3) objectives of learning analytics (learner and teacher support), (4) risks & challenges, (5) data management, and (6) research & data analysis. The rating results of the these statements show an obvious drop of rating scale in the '*ease of implementation*' level of these themes, compared to their '*importance*' level. One of the implications is that the six features could potentially be challenges to deal with in order to scale up the adoption of LA.

3.2 State of adoption – senior managers' perspectives

The interview data showed that 21 out of 51 institutions were already implementing centrally-supported learning analytics projects, 9 of which had reached institution-wide level, 7 partial-level (including pilot projects), and 5 were at the data exploration and cleaning stage. Meanwhile, 18 institutions were in preparation to roll out institutional learning analytics projects, and 12 did not have any concrete plans for an institutional learning analytics project yet.

The survey data revealed that 15 institutions had implemented learning analytics, of which 2 had reached full implementation and 13 were in small scale testing phases. Sixteen institutions were in preparation for learning analytics projects, and 15 were interested but had no concrete plans yet.

One of the implications of the two data sets is that there is high interest in LA among HEIs in Europe, but the maturity of adoption is low.

3.3 Top challenges associated with ROMA components

Our mapping of the institutional interviews identified key themes of challenges associated with each of the six components of the ROMA framework. Among these, two top challenges are methods used to implement LA and issues around ethics and privacy.

4 CONCLUSION

The SHEILA project has reached out to nearly half of the European countries, and observed high interest in LA among HEIs. However, few HEIs have taken a systematic approach to LA with defined strategy and policy. Our preliminary findings have identified prominent challenges that need to be tackled through an overarching policy.

REFERENCES

Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D. (2014). Embracing Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Challenge. *Research & Practice in Assessment*, *9*, 17–28.

Video Exemplar; From Meaningless to Meaningful

Holmesglen Institute
cedomir.gladovic@holmesglen.edu.au

ABSTRACT: Without industry experience, it is often difficult for students to understand the quality of work required in the discipline. The concept of the video exemplar was created as an educational tool to help students better understand the assessment criteria and to define quality. The exemplar uses two examples of the student's work and one example of an industry report and combines them into one comprehensive document. The video exemplar uses computer software for the screen capture and audio narration. A webcam captures the facial expressions of the presenter to create an even more realistic environment of a presentation by a professional (tutor) and identifies important concepts, contents and quality of work from the exemplar. The objectives of this pilot were to help students better understand the assessment requirements and to identify the explicit quality standards they should achieve when creating their own reports. This video exemplar also triggered proposal of the analytics dashboard that will provide students with comparison analysis based on the exemplar content, the student's up to date work and work of other students undertaking the same project.

Keywords: Video exemplar, analytics, feedback, assessment, learning.

1.1 Introduction

Numbers: The proposed poster aims to use a video exemplar as a method to enhance learning experiences for students from the *(name of the course)* at *(name of educational institution)*. Carless and Chan (2016) defined exemplar as 'Carefully chosen samples of student work which are used to illustrate dimensions of quality and clarify assessment expectations'. The proposed pilot encapsulated not only samples of students' work but also samples of reports and case studies from the industry. Our assumption as educators is that by providing suitable exemplars, we will help students achieve their assessment goals and also enable them to define quality. One of the initials steps conducted by students when exemplars are provided is closely replicating or mimicking the structure of exemplars. Under such circumstances, however, exemplars may become 'disablers' rather than 'enablers of learning'. Another issue that was evident among BEDP students was feeling overwhelmed at the experience due to the length and richness of the content provided in the form of exemplars. To address such issues the idea of a video exemplar was initiated.

Students' abilities to monitor, evaluate and regulate learning (Ajjawi & Boud, 2017) is one of the main educational objectives. We hold a view of the BEDP Department that exemplars are an additional form of formative feedback with the intention to help students with their evaluative judgement capabilities. The ability of students to identify quality in the work of others is also one of the most desirable professional characteristics. Video exemplars will help them with their ability to judge their own work and to define and achieve the required quality.

1.2 Video Tools as Learning Tools

The use of videos in an educational context is not a new phenomenon. Henderson and Phillips (2015) used video-based feedback as an alternative to text-based feedback, while Crook et al. (2012) argued that screen-capture technology in the form of a video had potentials to support learning differently from other technologies. Such a view is also supported by Cann (2007) who declared the podcast as dead and introduced video technology as a more engaging tool to support learning. These studies were drivers toward the idea that providing exemplars with or without text annotations would not be as sufficient as providing students with an exemplar in the video format.

1.2.1 Video Exemplar (Pilot)

The video exemplar has been developed for a second-year subject to help students use their own evaluative judgement capabilities to reflect and evaluate the quality of others' work critically. We sampled two examples of the student's work from previous semesters and one example of an industry report, combining them into one comprehensive document. The video exemplar uses computer software for screen capture and audio narration. The webcam captures the tutor's facial expressions. Our actions, as we moved throughout the electronic document using audio narration to outline the important concepts and content of the document, are video recorded (Figure 1).

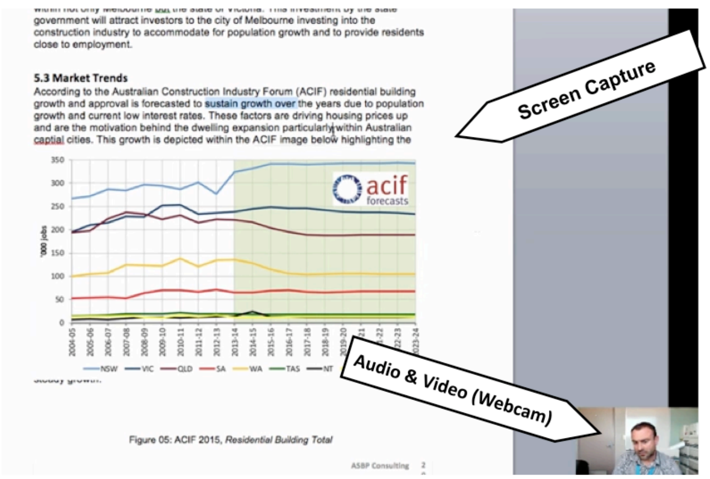


Figure 1: Example of video-exemplar (extract from video)

The intention was to inform students about project requirements and what the explicit quality standards they should achieve when creating their own reports, as they often find it difficult to understand the quality of work required in the discipline.

1.2.2 Analytic Dashboard

The added value of video exemplar would also be evident in the proposed use of Analytic Dashboard tool. The analytic dashboard will help students enter data to compare: a) project requirements; b) their own up-to-date project progress; and c) exemplars outcomes. Such data, once shared on the analytic dashboard, will be visible to other students allowing the community of learners to identify gaps in their own knowledge and to improve it accordingly.

The conceptual idea of the Analytic Dashboard Tool has been illustrated in figure 2.

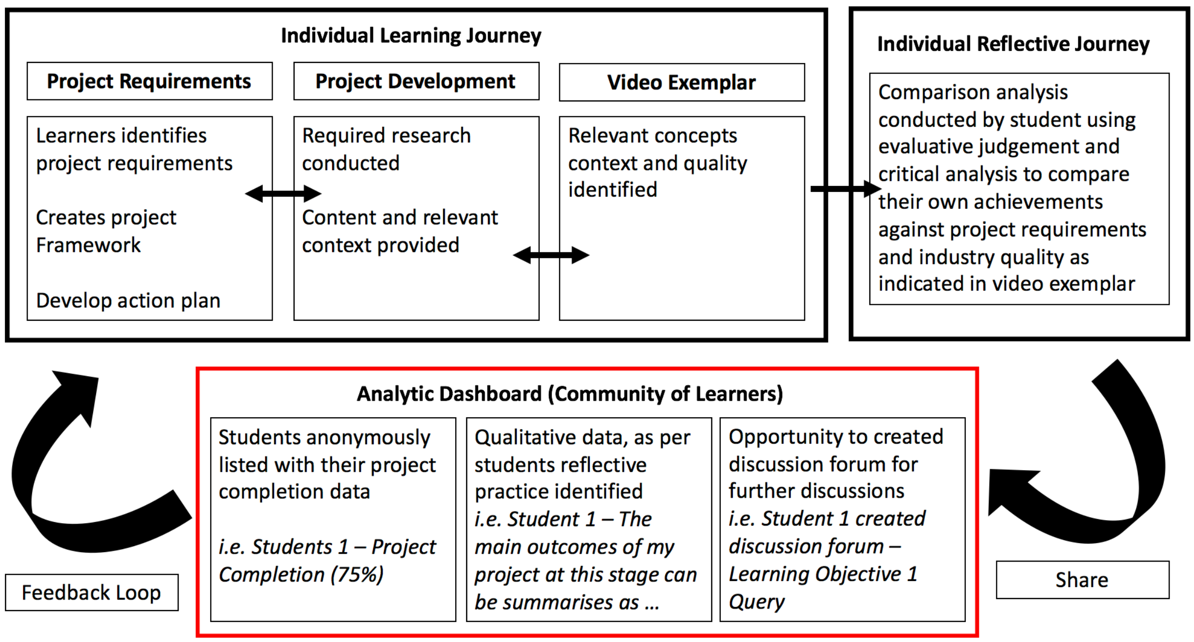


Figure 2: Analytic dashboard tool

The analytic dashboard will enable learners and their assessor to evaluate their own progress, quantitatively and qualitatively, using learning analytics as a feedback dialogue.

REFERENCES

Ajjawi, R., & Boud, D. (2017). Researching feedback dialogue: an interactional analysis approach. *Assessment & Evaluation in Higher Education, 42*(2), 252-265. doi:10.1080/02602938.2015.1102863

Cann, A. J. (2007). Podcasting is Dead. Long Live Video! *Bioscience Education, 10*(1), 1-4.

Carless, D., & Chan, K. K. H. (2016). Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education*, 1-12. doi:10.1080/02602938.2016.1211246

Crook, A., Mauchline, A., Maw, S., Lawson, C., Drinkwater, R., Lundqvist, K., . . . Park, J. (2012). The use of video technology for providing feedback to students: Can it enhance the feedback experience for staff and students? *Computers & Education, 58*, 386-396.

Henderson, M., & Phillips, M. (2015). Video-based feedback on student assessment: Scarily personal. *Australasian Journal of Educational Technology, 31*(1), 51-66.

How am I Doing?: Student-Facing Performance Dashboards in Higher Education

Carl C. Haynes[1], Stephanie D. Teasley[1], Stephanie Hayley[2], Meghan Oster[3] & John Whitmer[4]
[1]School of Information, [2]Academic Innovation, [3]School of Education
University of Michigan 105 S. State St., Ann Arbor, MI 48103 USA
[4]Blackboard, Inc. 58 Maiden Lane, San Francisco, CA 94108 USA
[cchaynes, steasley, swooten, omeghan]@umich.edu, john.whitmer@blackboard.com

ABSTRACT: Interest in implementing student-facing performance dashboards in higher education is increasing, but there are few studies directly addressing how students perceive and interpret these dashboards. We presented 47 undergraduate students with simulated activity and performance feedback from a Learning Management System dashboard representing 3 time points in an academic semester (early, mid, late). Using a 2 (Feedback Condition: High, Low Performance) x 2 (GPA: High, Low) design, we investigated how students interpret this information, looking specifically at whether the content of the feedback and its consistency with students' overall academic achievement affected students' responses to the dashboard. Our results showed that most students found the dashboard visualizations informative. There were differences between students' potential use of such systems and how they interpreted the impact of the information provided based on their prior academic achievement and the specific feedback provided by the system. Low-GPA students, in particular, found the dashboard messages more useful than students with high-GPAs.

Keywords: Learning Analytics; Higher Education; Academic Technology; Dashboards; Feedback; Student Motivation.

1 INTRODUCTION

Learning analytics dashboards provide a powerful means to present information to different stakeholders in academia (Bodily & Verbert, 2017). Most existing educational dashboards are aimed at academic professionals such as administrators, advisors, and instructors; however, increasingly the users are students (Teasley, 2017). Student-facing dashboards have typically included features that display comparative feedback about students' performance relative to course peers (Beheshitha et al., 2016) intended to produce "actionable insight" (Broos et al., 2017). However, little is known about the impact of student-facing dashboards. Are students able to interpret the information provided by such systems, and do they know what to do with it? Perhaps most importantly, which students find this information motivating versus demotivating, and under which circumstances? We designed a study to investigate the following research questions based on those raised by the existing literature:

RQ1: Do students value dashboard notifications providing information about their academic standing in their courses?

RQ2: Do students value visualizations that present information about their course performance relative to their peers?

RQ3: Do students find comparative performance information motivating or demotivating, and does this vary by the nature of the feedback delivered (high vs. low performance) and academic standing (GPA)?

2 METHOD

Forty-seven undergraduates from a large research university in the United States participated in this study. Students were randomly assigned to one of two groups, a "high" performance feedback condition and a "low" performance feedback condition where they viewed a summary message and graphic presentation of their login activity and performance (grade summary) representing three time points in a semester (early, mid late). A survey administered after the session captured students' views (scale from 1 = Strongly Disagree to 5 = Strong Agree) about the dashboard content and students' general preferences for future use of this system.

3 RESULTS

Using a 2 (Performance condition) x 2 (GPA) design, we ran ANOVAs to test for significant differences between experimental conditions. The results are shown in Table 2 on the poster.

RQ1: *Do students value dashboard notifications providing information about their academic standing in their courses?* Students' opinions about the system were generally high, with no survey questions rated below M = 3.64. Two of the three highest overall ratings were for questions related to future use (Q10) M = 4.55 (range 4.20-4.70) and finding the graphs informative (Q7) M = 4.55 (range 4.43-4.80). Overall, the students preferred the graphs (Qs5-8) to the activity stream messages (Qs1-4).

RQ2: *Do students value visualizations that present information about their course performance relative to their peers?* Question 5, which asked students whether the graphs helped them understand their position in the course, received the highest overall rating of all the survey questions M = 4.68 (range 4.28-4.93).

RQ3: *Do students find comparative performance information (de)motivating, and does this vary by the nature of the feedback delivered (high vs. low performance) and academic standing (GPA)?* The only statistically significant differences were due to GPA for the two questions regarding activity messages. High GPA students in both feedback conditions were less motivated to take immediate action (Q1: F (2, 43) = 6.12, p < .02), less likely to turn on the summary messages (Q2: F (2, 44) = 5.31, p < .03) and would check them less often (Q3: F (2, 42) = 4.94, p < .03) than those with a Low GPA. Students in the Low-Performance condition were significantly more likely than students in the High-Performance conditions to find the follow-up actions useful (Q4: F (2, 44) = 6.07, p < .02).

REFERENCES

Beheshitha, S. S., Hatala, M., Gašević, D., & Joksimović, S. (2016, April). The role of achievement goal orientations when studying effect of learning analytics visualizations. In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (pp. 54-63). ACM.

Bodily, R., & Verbert, K. (2017, March). Trends and issues in student-facing learning analytics reporting systems research. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (pp. 309-318). ACM.

Broos, T., Peeters, L., Verbert, K., Van Soom, C., Langie, G., & De Laet, T. (2017, July). Dashboard for Actionable Feedback on Learning Skills: Scalability and Usefulness. In International Conference on Learning and Collaboration Technologies (pp. 229-241). Springer, Cham.

Teasley, S. D. (2017). Student Facing Dashboards: One Size Fits All?. Technology, Knowledge and Learning, 1-8.

Automated Extraction of Learning Goals and Objectives from Syllabi using LDA and Neural Nets

Charles Lang[Ŧ], Rinat Levy-Cohen[Ł], Charlotte Woo[Ŧ], Benjamin Roberts[Ŧ], Samantha Pepe[Ŧ], Roshan Verma[Ŧ] & Yan Xu[Ŧ]

Ŧ - Teachers College Columbia University

Ł - Fordham University

ABSTRACT: Effective application of learning analytics to the classroom requires the generation of reliable standards on which to base metrics (Ferguson, 2012). An effective, but highly labor intensive, way to create these standards is for instructional designers or instructors to define competencies, learning goals and objectives for their courses. Although this may be the ideal solution, it may be unrealistic to scale the approach, especially for courses that have already been designed and taught for many years. There is therefore a need for technological tools that can take the friction out of this process. Here we describe a process for leveraging syllabi that do not have learning standards defined to automatically extract learning goals and objectives. The process utilizes Latent Dirichlet Allocation (LDA) topic modelling coupled with neural nets to generate sentences describing learning goals and objectives directly from a syllabi for review by an instructor or course designer.

Keywords: competency based education, learning goals, Latent Dirichlet Allocation (LDA), topic modelling, Bloom's taxonomy, text imputation, neural net.

1 INTRODUCTION

Competency based learning is a model characterized by explicitly defined knowledge and skills, expediency in results and close alignment to industry demands (Voorhees, 2001). Tightly-defined learning goals and objectives become pertinent in ensuring that competency based learning achieves the intended outcome. Good learning goals and objectives are difficult and time consuming to generate (Piskurich, 2015). Any effort to automate the process of creating goals and objectives without compromising their quality will be valuable to educators and designers and crucial for feeding into analytic systems. This work builds on Ramesh, Sasikumar, & Iyer (2016) who developed an automated system to integrate course content and levels of thinking (as defined by the syllabus) into a learning objective annotated ontology (LAO) and connects to other efforts to aid instructional design such as Avila et al (2017) and Greer et al. (2015).

2 METHODS

Seventy syllabi were collected using convenience sampling. Variables from the headings were extracted as learning goals and those from the body as learning objectives using the PDFBOX and Cloud java libraries. Text was processed and latent Dirichlet Allocation (LDA) was performed to extract topics from this corpus using the tm, Snowball, NLP and tidytext R packages. Topics were then defined as the most common word per topic (Hsiao & Awasthi, 2015). The topic names were then manually matched to objective verbs from Bloom's hierarchy of learning objectives (1956). A three-layer neural net was then trained to predict which learning objectives matched with the generated topics using the neuralnet R package. The trained neural net was tested on a sample of three syllabi.

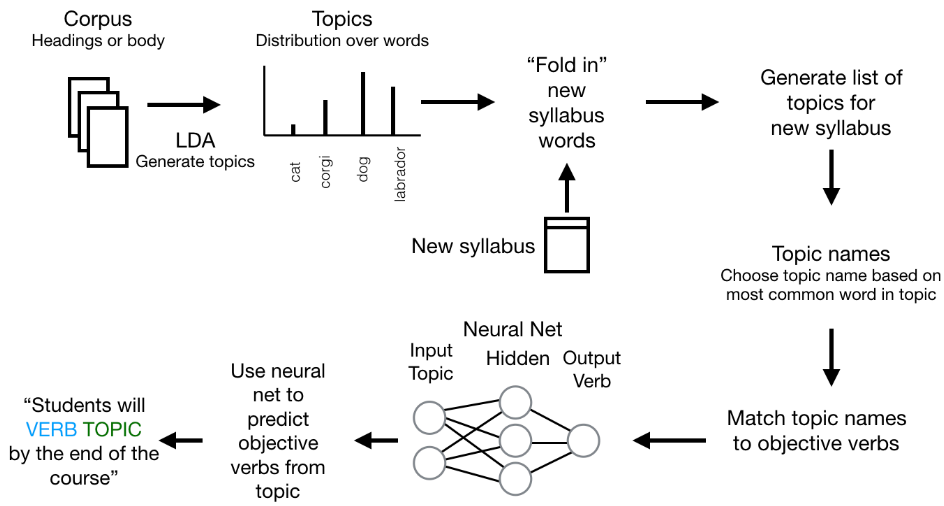


Figure 1: LDA and neural net process flow

3 RESULTS & DISCUSSION

In the future we hope to be able to create metrics on the efficacy of the system based on whether or not the sentences make sense to the syllabus author. We hope to create a system that we can make available to instructors and capture this information. We further hope there is utility for the learning analytics community in using such systems to aid the creation of useful standards to feed learning analytic systems.

REFERENCES

Bloom, B. S. (1956). *Taxonomy of Educational Objectives*. Vol. 1. New York: McKay, 20-24.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5–6), 304–317.

Hsiao, I. H., & Awasthi, P. (2015). Topic facet modeling: semantic visual analytics for online discussion forums. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge.* (pp. 231-235). ACM.

Piskurich, G. M. (2015). *Rapid instructional design: Learning ID fast and right*. New York: Wiley & Sons.

Ramesh, R., Sasikumar, M., & Iyer, S. (2016, July). Integrating the learning objectives and syllabus into a domain ontology for data structures course. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 266-271). ACM.

Voorhees, R. A. (2001). Competency-Based learning models: A necessary future. *New directions for Institutional Research*, 2001(110), 5-13.

Greer, J., Molinaro, M., Ochoa, X., & McKay, T. (2016, April). Learning analytics for curriculum and program quality improvement (2016). In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 494-495). ACM.

Avila, C., Baldiris, S., Fabregat, R., & Graf, S. (2017, March). ATCE: an analytics tool to trace the creation and evaluation of inclusive and accessible open educational resources. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 183-187). ACM

Identifying Trends in Student Success in Online Economics Courses at Colorado State University

Karen Gebhardt
Colorado State University
Fort Collins, CO USA
KarenG@colostate.edu

Kelly McKenna
Colorado State University
Fort Collins, CO USA
Kelly.McKenna@colostate.edu

Melanie G. Long
Colorado State University
Fort Collins, CO USA
MGLong@colostate.edu

ABSTRACT: Online learning has become increasingly important both as a way for students to access higher education and as a source of revenue for universities. Despite careful integration of best practices in online course design, many students are not as successful in online courses as compared to on campus courses. This poster presents an analysis of student demographic and course outcomes for more than 29,000 students taking online and on campus economic courses between 2012 and 2016 at Colorado State University. Results show that on campus students tend to be more successful as measured as the proportion of A, B, and C grades earned. Results related to student subgroups show that the differences in success rates for female students and ethnically diverse students are narrowed in online courses. These results indicate important heterogeneity when comparing patterns of success between students taking online and on campus courses and among students taking online courses. Careful consideration of such differences and further use of learning analytics can help inform effective online course interventions to improve student success.

Keywords: learning analytics, online learning, economics education

1 INTRODUCTION

Since 2005, the Economics Department at Colorado State University (CSU) has offered a broad selection of online courses. These courses have experienced a steady growth in enrollment where the number of enrollments has increased from 609 in 2012 to 977 in 2016, a 60.4% increase over the 5-year period far outpacing growth in enrollments in economics courses offered on campus (8.8%), economics majors (15.4%), and growth at the university (9.4%). All economics courses at CSU integrate best practices in online course design and are often taught by the same instructor teaching the same course on campus during the same term. However, despite the careful integration of best practices, many students are not successful in these online courses. Between 2012 and 2016, there has been a striking and consistent difference in success rates (i.e., the proportion of students receiving A's, B's, or C's) between the online courses (74.33%) and on campus courses (87.08%) (see Table 1).

Table 1: Success Rates by Delivery Mode and Term, Calendar Years 2012 - 2016

Delivery Mode	Spring		Summer		Fall		OVERALL	
	Enrolled	*Success Rate*	*Enrolled*	*Success Rate*	*Enrolled*	*Success Rate*	*Enrolled*	*Success Rate*
Online	1,057	70.96%	1,723	78.18%	983	71.21%	3,763	74.33%
On campus	11,999	87.04%	1,418	91.96%	12,259	86.55%	25,676	87.08%
Overall	13,056	85.74%	3,141	84.40%	13,242	85.41%	29,439	85.45%

This study seeks to understand this concerning differential in success when students take online or on campus courses. Careful consideration of such differences and further use of learning analytics can help inform effective online course interventions to improve success for students making use of this increasingly important delivery mode (Johnson, 2012).

2 METHODS

The purpose of this research is twofold: (1) to identify the difference in success rates along select parameters between online and on campus courses, and (2) to investigate the difference in demographic and academic characteristics among students who were successful and not successful in online courses. This study used demographic and course outcome data for students taking online and on campus economics courses at a large state university between spring 2012 and fall 2016. Only courses taught both online and on campus during the study period were included in the analysis. There were 29,439 student enrollments in these courses, of which 3,763 (12.8%) were online and 25,676 (87.2%) were on campus. Two statistical methods, cross-tabulations as well as regression analyses with course-term-delivery mode fixed effects and student-clustered standard errors, were used to explore the patterns of success as identified in the purpose of this research.

3 RESULTS

Regression analyses results point to beginning of semester GPA as the most consistent predictor of student success, but additional results show that significant heterogeneity in success exists between students taking online and on campus courses and among students taking online courses.

When comparing outcomes between students taking online and on campus courses, the data indicates the average success rate for students in online courses was less than that of students in on campus courses across all terms (i.e., spring, summer, fall). Results show that students taking online courses were more ethnically diverse, older, and less likely to be first-generation than their on campus counterparts. Female students earn higher grades than male students regardless of delivery mode but were less likely to be successful on average in online courses than on campus courses. In contrast, ethnically diverse students earn lower grades than non-ethnically diverse students regardless of delivery mode, but this difference was less pronounced among students taking online classes that are closely affiliated with the university. Success differentials between online and on campus courses are found to be smaller on average for lower division than upper division courses.

When comparing outcomes among students taking online courses, results show that generally, students who take a mix of online and on campus courses have the highest success rates among students taking online classes. More specifically, first-generation students who take a mix of online and on campus courses have higher success rates in online courses than non-first-generation students. Students who are closely affiliated with the university (e.g., degree seeking at this institution) have higher success rates in online classes than students who are not closely affiliated with the university (e.g., students from other universities).

REFERENCES

Johnson, N. (2012). *The Institutional Costs of Student Attrition*. Delta Cost Project at American Institutes for Research. Retrieved from https://eric.ed.gov/?id=ED536126

Estimation of Learners' Programming Level based on Free Descriptions in MOOC Course

Masako Furukawa, Kazutsuna Yamaji
National Institute of Informatics
{furukawa, yamaji}@nii.ac.jp

ABSTRACT: National Institute of Informatics released a basic course for software programming by MOOC. In such online courses, it is important to support learners according to their levels and needs. To realize such support for learners, we propose to estimate the learners' programming level based on free descriptions in MOOC course. The difference of used keywords for each programming level is revealed by the correspondence analysis.

Keywords: MOOC, Free descriptions, Visualization, Recommendation

1 INTRODUCTION

In online courses, it is important to support learners according to their levels and needs. For example, Sun et al. (2008) clarified the critical factors influencing learner satisfaction. To realize such support for learners, we propose to estimate the learners' programming level based on free descriptions in MOOC course.

2 METHODS

National Institute of Informatics released a basic course for learning programming by MOOC. The lecture took place from August 9, 2016, and lasted 4 weeks. The number of learners was 6,859. Each lecture consists of 3 to 5 videos and a confirmation quiz. We set up a discussion board for discussion and mutual help. The number of discussion threads reached 210. Before starting the course, the learners answered questionnaires. One of the questionnaires asks the learners their programming levels as shown in Table 1. The number of respondents who answered the questionnaires and the number of contributors to the discussion board are also shown. To clarify the relationship between programming level and free descriptions in discussion board, we use correspondence analysis.

Table 1: Programming levels and the number of messages / contributors

ID	*Programming level*	*Number of respondents*	*Number of contributors*
A1	I have no programming experience	1,313	58
A2	I have studied programming by primary books or sites	601	35
A3	I can make a program	392	34
A4	I program daily	119	6

ID	Programming level	Number of respondents	Number of contributors
NA	(No response)	4,053	16

The correspondence analysis is a method used in the natural language analysis, which visualizes relationships among categories and keywords. At first, 30 keywords were extracted in order of appearance count. Then, the frequency of these keywords is calculated for each programming level. The vector of frequency of keywords is reduced to two-dimensional using correspondence analysis.

3 RESULTS

Figure 1 shows the results of correspondence analysis. The labels of "A1" - "A4" and "NA" correspond to the IDs in Table 1. The circles show the location and frequency of the keywords. When we see the location of "A2", there are many segments of program codes. This shows that the learners who have studied programming by primary books or sites frequently wrote concrete program codes in discussion board. When we see the locations of "A3" and "NA", there are keywords such as "File" and "Character". This is because the learners who say "I can make a program" tried to explain about manipulation of files and characters. The locations of "A1" and "A4" are similar, and the keywords "Execute" and "Error" are around there. This is because they discussed errors of programs. "A1" frequently used "Thank" to express appreciation.

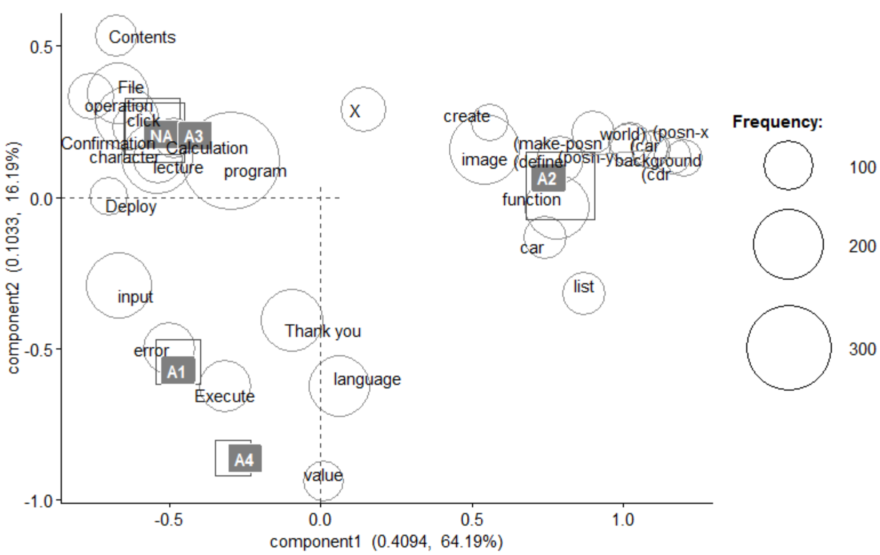


Figure 1: Results of correspondence analysis

4 CONCLUSION

In this paper, we proposed to estimate the learners' programming level based on free descriptions in MOOC course. By the correspondence analysis of discussion board, the difference of used keywords for each programming level was revealed.

REFERENCES

Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y. and Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. Computers & Education, Volume 50, Issue 4, 1183–1202.

Estimating Grades from Students' Behaviors of Programming Exercises Using Machine Learning

Toshiyasu Kato
Nippon Institute of Technology, Japan
katoto@nit.ac.jp

Yuki Terawaki
Hollywood Graduate School of Beauty Business, Japan
terawaki@hollywood.ac.jp

Yasushi Kambayashi
Nippon Institute of Technology, Japan
yasushi@nit.ac.jp

Yasushi Kodama
Hosei University, Japan
yass@hosei.ac.jp

ABSTRACT: The prediction of students' final grades in an early stage is an important task in the field of learning analytics. Programming exercises are time-consuming activities, and individual students' programming behaviors are quite different from each other's, even when they are solving the same problem. We conjecture that the final grade can be estimated from programming behaviors. We estimated the grade of 2016 using the data of 2014 and 2015. From this result, we confirmed that the grade can be estimated by analyzing programming behaviors.

Keywords: Programming Exercises, Teaching Assistants, Learning analytics, Deep Learning, Grading estimate

1 INTRODUCTION

By analyzing the logs of students using data mining techniques, we can determine learning patterns of students, which helps teachers in detecting "at-risk" students (Baradwaj, 2011). During programming exercises at institutions of higher education, students individually tackle assigned problems. Some students can easily solve problems, while many other students spend too much time on single problems. It is thus important to identify in an early stage which students are likely to have such programming difficulties.

The prediction of students' final grades in an early stage is an important task in the field of learning analytics, e.g., investigated in using regression analysis (You, 2016). In order for a teacher to grasp the timing when he or she intervene students who need help, this paper estimates the students' grades. The authors have developed and used an unique programming exercise support system (called PRESS) to identify the programming behaviors discussed herein (Kato, 2012).

2 ANALYSIS OF PROGRAMMING BEHAVIORS AND GRADES

In this paper, we propose a method for predicting students' final grades by a neural network approach, using the programming behaviors of the PRESS system. In this study, we prepared data on the programming behaviors of students in 2014, 2015 and 2016 classes. Table 1 shows the programming behavior data and final grades for 2014, 2015 and 2016 for each explanatory variable. These data are average values (excluding the number of students and grade). The grade is determined by the final paper test.

Table 1: Programming behaviors and grades.

	2014	2015	2016
Number of students	85	80	107
Problem-solving time [s]	815	936	639
Average compile interval [s]	295	352	405
Average execution interval [s]	513	584	139
Compile times	10	8	7
Number of errors until problem resolution	7	3	5
Number of errors of the same type that are most frequent	4	3	2
Number of students with long problem-solving times	32	35	51
Number of problem resolutions	40	41	38
Grade 1: Failing	17	1	5
2: Passing	33	13	34
3: Average	17	17	34
4: Good	21	36	30
5: Excellent	2	15	7

We estimated the grade of 2016 using the data of 2014 and 2015. The results of various data mining are shown in Table 2. From this result, we confirmed that the grade can be estimated by analyzing programming behaviors. Moreover, it was confirmed that high accuracy is obtained by using deep learning. Our method can find students who are going to fail from the class.

Table 2: Estimation results.

	Accuracy	Precision	Recall
Deep Learning	0.70	0.41	0.46
Random Forest	0.62	0.34	0.37
SVM	0.58	0.32	0.42
Naive Bayesian	0.58	0.32	0.38

ACKNOWLEDGMENT

This work was supported by Japan Society for Promotion of Science (JSPS), with the basic research program (C) (15K01094), and (A) 26245048 and 26240008, Grant-in-Aid for Scientific Research.

REFERENCES

Baradwaj, B. & Pal, S. (2011). Mining Educational Data to Analyze Student's Performance, *International Journal of Advanced Computer Science and Applications*, 6, 2, 63-69.

Kato T., Ishikawa T. (2012). Design and Evaluation of Support Functions of Course Management Systems for Assessing Learning Conditions in Programming Practicums. *ICALT 2012*, 205–207.

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning, *Internet and Higher Education*, 29, 23-30.

Report from the Learning Analytics Summit in Medical Education at ICRE 2017

Author(s): Dr. Rodrigo Cavalcanti; Dr. Jason Frank; Dr. Stan Hamstra; Dr. Karen Hauer; Dr. Martin Pusic; Dr. Tim Shaw; Dr. Brent Thoma; Dr. Amol Verma; and, Dr. Eric Warm.
For the Royal College of Physicians and Surgeons of Canada
Rodrigo.cavalcanti@uhn.ca

ABSTRACT: Learning analytics shows great promise in medical education as the field transitions towards a competency-based approach. Learning analytics can benefit individual learners and tutors, and can also be applied to promote programmatic assessment at the training program level. Furthermore, at a national specialty level, data-comparisons across programs can help inform standards of training. Areas requiring development for successful implementation of learning analytics in medical education include triangulation of different data-sources needed to document of specific competencies, as well as careful consideration of data stewardship, privacy and legacy.

Keywords: Medical Education; Programmatic assessment; Competency-based education

1 THE LEARNING ANALYTICS SUMMIT IN MEDICAL EDUCATION

This paper provides and overview of a two-day Summit on Learning Analytics in Medical Education held prior to the International Conference in Medical Education in Quebec City in Oct 2017. The summit was designed to bring together a network of medical educators interested in discussing learning analytics applications. It summarizes some of the key discussion topics and highlights potential applications and considerations when extending learning analytics tools to a professional training context such as medical education.

1.1 Benefits of learning analytics at three levels: learner to program to specialty

Learning analytics is a growing field in higher education that aims to implement tools to help in "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens 2011). Learning analytics can help inform progression and learning both through feedback on previous activities and in helping design future learning opportunities. Similarly, in medical education analytics tools also show great promise to help inform growth in clinical knowledge and skills **(Cirigliano et al., 2017; Ellaway, Pusic, Galbraith, & Cameron, 2014)**. Increased assessment data from evolution towards Competency-Based Medical Education (CBME **(Frank et al., 2010)**) creates opportunity to leverage the increased volume of assessment data. Learning analytics tools can be applied at three levels: to help trainees and supervisors in structuring clinical learning; to benefit programs in understanding progression of trainees; and to inform specialties at the national level regarding standards of training.

Applications at the learner level can help collate performance data from across learning experiences and supervisors and present it synoptically to help gauge progression and inform areas for future growth. When paired with competency frameworks (Frank et al., 2010), (Iobst et al., 2010)), these tools can provide powerful ways to make connections between skills acquired in different contexts, and can provide unique platforms for coaching learners.

At the program level, aggregate performance data across learners can provide invaluable support for programmatic assessment (Dijkstra, van der Vleuten, & Schuwirth, 2009), highlighting areas in need of curricular enhancement. Data can also be analyzed at the clinical rotation level to provide impetus and direction for faculty development initiatives (Warm).

Similar to program level analyses, learning analytics tools that aggregate data across multiple training programs for a specific specialty at a national level indicate how objectives of training are being met (Conforti et al., 2017). By providing a national overview, analyses of aggregate learning curves can inform decisions regarding duration of training, breath of required training experiences and best timing for acquiring highly-specialized skills.

1.2 Data sources for analytics in medical education

Learning analytics in higher education draw data from online learning activities, assessments, and demographic data. The workplace-based nature of medical training allows incorporation of additional sources of data into learning analytics tools, including workplace-based assessment, and clinical practice data.

In addition to more established sources of data used for learning analytics, such as demographics, online learning activities, and written assessment data, medical education, particularly at the postgraduate level, offers some unique data sources. These include other assessment formats such as objective-structured clinical assessments (OSCEs - (Shumway, Harden, Association for Medical Education in Europe, 2003)), a multi-station oral exam designed to assess skills and knowledge required to perform in a clinical setting. More unique types of assessment data in medical training include direct observation of clinical skills through workplace-based observations by tutors, as well as high fidelity simulation-based assessments, and performance data derived from clinical activities, such as resource-utilization (e.g. number and types of tests ordered), clinical parameters (such as achievement of target parameters in blood pressure or cholesterol profile), and patient outcomes.

At the postgraduate level, competency frameworks such as CanMEDS, or the ACGME competencies can also provide a taxonomy that allows learning and assessment data from different contexts to be organized into a overview of skills acquisition. Lastly, across the continuum from medical school, through residency, and into continuing education may allow the development of a lifetime portfolio of learning to be available for practitioners.

1.3 Data stewardship: considerations in medical education

In any learning analytics implementation, questions regarding data stewardship are important to consider. Concerns regarding access to the data raise questions of privacy. In a professional training setting such as medical education these concerns are amplified by potential medico-legal ramifications. These include issues of access to private health information, as well as potential litigation for the practitioners based on documented past performance. These require very careful design, with clear safeguards for learners, so as to ensure that the main objective of learning analytics tools remains the improvement of learner performance in a safe and effective manner. Paramount to these concerns is a clear delineation of who has access to learner data in which capacity, and which learner-centered goals are achieved by granting access.

Issues of data-legacy arise as a concern in medical education as practitioners pass through stages of training and certification. First, development of valid mechanisms, based on learning analytics, will be required to define progression through training. Secondly, individuals may want information on formative experiences eliminated from their records once they are certified for practice. Discussions regarding how to handle these issues are just beginning in medical education, but it will be crucial for educational institutions and governing bodies to clarify prior to implementation.

In conclusion, learning analytics is expanding into medical education across the continuum from undergraduate, to postgraduate training. These tools show great promise to inform growth in the filed from helping individual learners to setting specialty standards. However, their development will require integration of multiple novel data sources and careful consideration of many issues in data-stewardship.

REFERENCES:

Cirigliano, M. M., Guthrie, C., Pusic, M. V., Cianciolo, A. T., Lim-Dunham, J. E., Spickard, A., & Terry, V. (2017). “Yes, and …” Exploring the Future of Learning Analytics in Medical Education. *Teaching and*

Learning in Medicine, *29*(4), 368–372. http://doi.org/10.1080/10401334.2017.1384731

Conforti, L. N., Yaghmour, N. A., Hamstra, S. J., Holmboe, E. S., Kennedy, B., Liu, J. J., et al. (2017). The Effect and Use of Milestones in the Assessment of Neurological Surgery Residents and Residency Programs. *Journal of Surgical Education*. http://doi.org/10.1016/j.jsurg.2017.06.001

Dijkstra, J., van der Vleuten, C. P. M., & Schuwirth, L. W. T. (2009). A new framework for designing programmes of assessment. *Advances in Health Sciences Education : Theory and Practice*. http://doi.org/10.1007/s10459-009-9205-z

Ellaway, R. H., Pusic, M. V., Galbraith, R. M., & Cameron, T. (2014). Developing the role of big data and analytics in health professional education. *Medical Teacher*, *36*(3), 216–222. http://doi.org/10.3109/0142159X.2014.874553

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., et al. (2010). Competency-based medical education: theory to practice. *Medical Teacher*, *32*(8), 638–645. http://doi.org/10.3109/0142159X.2010.501190

Iobst, W. F., Sherbino, J., Cate, O. T., Richardson, D. L., Dath, D., Swing, S. R., et al. (2010). Competency-based medical education in postgraduate medical education. *Medical Teacher*, *32*(8), 651–656. http://doi.org/10.3109/0142159X.2010.500709

Shumway, J. M., Harden, R. M., Association for Medical Education in Europe. (2003, November). AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Medical Teacher*. http://doi.org/10.1080/0142159032000151907

Testing temporal hypotheses about response behavior in Q&A forums using a statistical network model and meta-analysis

Chelsea Daniels **Yoav Bergner** **Collin Lynch** **Tiffany Barnes**
New York University North Carolina State University
yoav.bergner@nyu.edu

ABSTRACT: In the e-learning context, social network analysis (SNA) can be used to build understanding around the ways students participate and interact in online forums. This study contributes to the growing body of research that uses statistical methods to test hypotheses about structures in social networks. Specifically, we show how statistical analysis can be used to examine the effect of students' past online interactions on their later participatory behavior, and to synthesize results across multiple networks.

Keywords: MOOC, social network analysis, ERGM, meta-analysis

1 INTRODUCTION

With the emergence of massive open online courses (MOOCs) and increasing reliance on computer supported collaborative learning (CSCL) in traditional educational contexts there is an ongoing need to increase understanding around the technical and social interactions that define learning communities online [1, 2]. The application of statistical network models, such as exponential random graph models (ERGM), to SNA in CSCL provides a means by which to examine whether social network structure can be explained by hypothesized mechanisms [3]. This study builds upon recent applications of ERGM to CSCL [3, 5], and explores students' online behaviors in terms of the interactions that contribute to the structure of a social learning environment.

2 CONTEXT AND DATA

In this study we are interested in testing what we term the "pay-it-forward" hypothesis: whether having received a response in previous weeks will contribute to an individual responding to someone else in a given week. This paper looks at the pay-it-forward hypothesis in the context of two different learning environments: (I) a blended computer science course at NCSU with 251 registered students (DM), and (II) a Coursera MOOC with 48,000 total enrollees (BDE). A key contribution of the present work is in the treatment of each week's response network as a sample of behavior at a particular time slice. Results from multiple weeks are then combined by means of a meta-analysis (two-level model). By combining the networks in a multi-level model, we are able to perform a hypothesis test regarding the course as a whole.

3 METHODS AND RESULTS

To study forum participation, we constructed a series of networks where students and instructors are nodes, and responses become directed ties. Forum participation data was segmented by week, and the number of responses a user received in the preceding weeks was modeled as a node covariate using an ERGM to test our hypothesis in a given week.

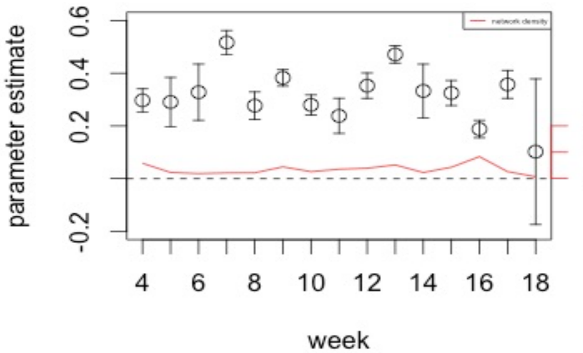


Figure 1: "Pay-it-forward" parameter estimates in DM dataset. Estimates begin in week 4 with a look-back window of three weeks.

Table 1: Two-stage least squares meta-analysis of "pay-it-forward" parameter estimates

	Meta-analytic effect size μ	Standard error
DM_course	0.336***	0.017
DM_wk6	0.367***	0.018
BDE_course	0.155***	0.020

*** indicates $p < 0.0001$

Our base model controls for density, degree distribution, asymmetry of ties, and instructor status. Previous replies are aggregated over a three-week window to account for the fading significance of past replies. In addition to reporting meta-analysis results for the duration of the course, we also perform a sensitivity analysis on the size of the look-back window. If the "pay-it-forward" hypothesis holds, we expect the a positive parameter estimate. Figure 1 shows a plot of the estimates by week. The estimated mean meta-analytic effect is shown in Table 1. Also shown are results for a particular week (DM_wk6) by varying the look-back window from 1-5 weeks, and results for the BDE course.

In sum, we show a statistical method for testing hypotheses about network behavior in online Q&A forums and a meta-analysis to temporal effects by means of a. It works for courses where student populations are stable as well as online courses with significant attrition effects.

REFERENCES

1. Haythornthwaite, C. (2001). Exploring multiplexity: Social network structures in a computer-supported distance learning class. *The information society*, *17*(3), 211-226.
2. Rabbany, R., ElAtia, S., Takaffoli, M., & Zaïane, O. R. (2014). Collaborative learning of students in online discussion forums: A social network analysis perspective. In *Educational data mining* (pp. 441-466). Springer International Publishing.
3. Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & De Kereki, I. F. (2016, April). Translating network position into performance: importance of centrality in different network configurations. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 314-323). ACM.
4. Zhu, M., Bergner, Y., Zhang, Y., Baker, R., Wang, Y., & Paquette, L. (2016, April). Longitudinal engagement, performance, and social connectivity: a MOOC case study using exponential random graph models. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 223-230). ACM.

Gender differences in confidence reports and in reactions to negative feedback within adaptive learning platforms

Ani Aghababyan[1], Nicholas Lewkow[1], Sean Burns[2], Jose Suárez-García[2]
[1]McGraw-Hill Education
[2]Colorado State University

ABSTRACT: Research shows that women typically express less confidence than men when evaluating their own performance and abilities. Moreover, some studies present evidence that compared to men women's confidence declines to a larger degree when they experience negative feedback. This study investigates whether men and women display similar levels of confidence and, when faced with negative feedback, experience similar declines in confidence in a safe-to-fail adaptive learning platform. Additionally, we explore whether both genders require a similar amount of time for confidence recovery post negative feedback (recovery time is defined as the number of questions before again reporting high confidence). This study captures data from 332 first- and second-year students who took one of three language courses offered at a major state university in the United States. While we find that men demonstrate significantly more overconfidence than women, the results show no significant gender differences with regard to negative feedback and confidence recovery times.

Keywords: confidence, gender bias, negative feedback, language learning, adaptive learning platform, formative feedback

1 INTRODUCTION

Considerable research has indicated that student's self-confidence is vital for their academic success (Adams & Ewen, 2009). Research has also suggested the existence of gender bias in a lot of aspects of one's life one of which is one's task-specific confidence (Yestrumskas, 2004). Additionally, research suggests that in circumstances where both women and men receive negative feedback (feedback that points out answer's incorrectness) on their performance, women's confidence suffers a much greater decline than men's (Roberts and Nolen-Hoeksema, 1989). It is expected that this perception of failure may create a more long-term effect on women's feeling of insecurity.

Nationally, women comprise over 57% of learners in public and private universities and over 60% of learners in language courses at the studied university who are using this learning platform. As such, it is vital to ensure that these systems are not designed to inadvertently hamper women's success or diminish their growth by gender biased instructional design. Therefore, in this study, we explore overall gender biases and potential differences in the reception of negative feedback in an adaptive learning platform. We hypothesize that the safe-to-fail environment the learning system provides creates an atmosphere for women to feel confident and diminishes gender difference.

2 DATA & CONTEXT

The studied university has a strong language department with Spanish, French, and German courses that are highly attended and offer several sections each semester. Our data set originates from 18 sections of 8 courses covering 3 foreign languages taught during the 2017 Spring semester. We analyzed data from 332 students enrolled in these courses.

To teach these courses, the university uses a combination of learning materials and platforms one of which is this safe-to-fail adaptive learning platform. For every assignment, students answered language questions and immediately reported their level of confidence in their answer on a four-level scale ranging from "no confidence" to "absolute confidence." There were 206 female students who answered 353,130 questions in total throughout the semester and 117 male students who answered 189,014 questions. There were twice as many questions answered by women compared to men.

3 ANALYSES

Independent-samples t-tests were conducted to compare average confidence, confidence variability (entropy), and over/underconfidence ratios between men and women. While our findings show insignificant differences between the genders on almost all measures, we find a significant difference exists in terms of overconfidence ratios among men (M=0.33, SD=0.31) and women (M=0.25, SD=0.30), t=2.3, p=0.01. This result suggests that when both women and men answer questions incorrectly, men are more likely to choose a higher confidence rating than women. This is consistent with previous research findings (Barber & Odean, 2001).

To compare the impact of negative feedback, we calculated the number of times students of either gender reduced their confidence immediately after receiving negative feedback that did not match their high confidence or expectations. Independent-samples t-test show that there was no significant difference between female and male students in the number of times negative feedback affected them. There was also no significant difference between the groups in terms of confidence recovery, which we define as the number of questions a student completes between receiving negative feedback and recovering their high confidence.

4. CONCLUSIONS & FUTURE WORK

Safe-to-fail adaptive learning platforms should provide a safe environment for learning through continuous practice. Showing no significant difference in the general confidence between men and women along with no difference in feedback effects suggests that this learning platform could be considered a safe place to learn by all users despite their gender identity. Additionally, showing that men tend to be more overconfident when answering incorrectly further reinforces previous research. In the future, we will revisit these analyses for women and men by studying both social science and physical/life science courses to factor in course difficulty and male-to-female enrollment ratio. Furthermore, we plan to incorporate measures of students grit and growth mindset in our gender difference analyses and split the groups by course modality (face-to-face vs. online).

REFERENCES

Adams, T. M., & Ewen, G. W. (2009, August). The importance of confidence in improving educational outcomes. In *25th annual conference on Distance Learning and Teaching*.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics*, *116*(1), 261-292.

Roberts, T. A., & Nolen-Hoeksema, S. (1989). Sex differences in reactions to evaluative feedback. *Sex Roles*, *21*(11-12), 725-747.

Yestrumskas, A. (2004). Should I choose the harder path? Gender differences in an experiment on task choice.

Examining the influence of socially shared metacognition on group problem solving

Author(s): Okan Arslan[1]; Wanli Xing[1]; Paul Howwitz[2]; Cynthia McIntyre[2]
Texas Tech University[1]; The Concord Consortium[2]
okan.arslan@ttu.edu; wanli.xing@ttu.edu; phorwitz@concord.org; cmcintyre@concord.org

ABSTRACT: The purpose of this study was to discover students' socially shared metacognition (SSM) patterns with a holistic way in a physics course supported by CSCL environment and to understand how SSM patterns influenced the task performance. To investigate patterns of SSM behaviors, 126 collaborative physics learning scenarios were manually coded. Then K-Means clustering analysis was conducted. As a result, four clusters were discovered for collaborative problem solving. Interaction between the four identified patterns with the final performance, using Chi-squared test, were examined. There was a significant difference on the distribution of success and failure groups among the clusters.

Keywords: CSCL, socially shared metacognition, physics course, problem solving

1 INTRODUCTION

Socially shared metacognition (SSM) is an important component of collaborative problem solving process by monitoring social level of cognitive process and collective memory, regulating external representation of tasks (Iiskala et al, 2011). Research studies reveal that collaborative learning is more effective to achieve high learning goals than individuals working alone (Järvelä, Hurme & Järvenoja, 2011). Previous studies have focused on isolated effects of these SSM dimensions (activate, confirm, change, and slow) on problem solving skills. It is also important to examine the combination of SSM dimensions in a holistic way to advance our understanding of SSM in collaborative learning. Thus, the purpose of this study is to discover students' SSM patterns with a holistic way in a physics problem based activity supported by CSCL environment and to understand how SSM patterns influence the team task performance.

2 METHOD

2.1 Design

Students were asked to solve problems about electronics circuits for Physics course. They used a computer supported collaborative system, named Teaching Teamwork Activities developed by The Concord Consortium. There were 5 different levels from easy to difficult one. Students were supposed to solve the problems gradually. Each problem activity was identified as an episode. They might deal with at least 1, at most 5 problem activities, episodes. Students needed to solve the problem in a certain time to be able to work with the next one.

2.2 Data Gathering and Analysis

Log Data were gathered from 30 groups consisted of 3 members by using Teaching Teamwork. In total there were 126 episodes to analyze. For K-means cluster analysis, the frequency of each SSM dimension were calculated. In terms of task performance, success and failure rate of the groups were calculated for the Chi-squared test of independence.

3 RESULTS

Cluster analysis was performed to achieve the greatest homogeneity in each group and the greatest differences among them to find out optimal number of cluster. According to Hartigan (1975)'s statistic, the optimal number of cluster was 4. Clusters were determined upon both SSM dimension frequencies and collaborative achievement status, namely, *active changers and socially high achievers*, *active confirmers and socially high achievers*, *active slowers and socially moderate achievers*, and *passive changers and socially moderate achievers*.

Further, the interaction between the four identified patterns with the final performance of the collaborative groups were examined by using Chi-squared test of independence. According to the results, there is a significant difference on the distribution of success and failure groups in each pattern (x2(3) = 8.72, p < .05). Mosaic graph with Pearson residuals (See Figure 1) was used to interpret the relationship between clusters and performance.

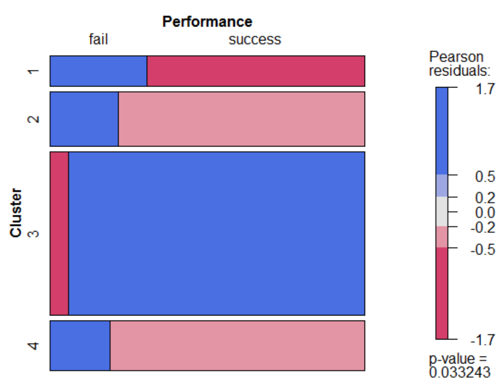


Figure 1: Mosaic Graph about relationship between clusters and performance

In terms of performance, *active confirmers and socially high achievers* tends to be the most successful group. *Active changers and socially high achievers* has the second most probability to be successful. *Passive changers and socially moderate achievers,* and *active slowers and socially moderate achievers* have the third and fourth most probability to be successful respectively. This significant result, to some extent, validates the discovered four patterns.

REFERENCES

Iiskala, T., Vauras, M., Lehtinen, E., & Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction,* 21, 379-393.

Järvelä, S., Hurme, T.-R., & Järvenoja, H. (2011). Self-regulation and motivation in CSCL environments In S. Ludvigsen, A. Lund & R. Säljö (Eds.), *Learning in social practices: ICT and new artifacts-transformation of social and cultural practices*. Pergamon.

Evaluation of Potential Impact of Transition to a BYOD Model via Exploratory Analysis of Device Usage Behaviors and Ownership Using a University's Network and Computing Resources Logs

Author: Béatrice Moissinac
Oregon State University
Beatrice.moissinac@oregonstate.edu

Author: Robin Pappas
Oregon State University
Robin.pappas@oregonstate.edu

ABSTRACT: Poster. A middle-sized 4-year United States University is considering transitioning from supporting student computing via computer labs and loaned laptops to a Bring Your Own Device (BYOD) model. BYOD models are supported by recent studies from the Pew Research Center and ECAR, indicating that between 88 and 90% of undergraduate students own a laptop. However, these studies rely on interviews and surveys, which are inherently limited to students' self-reports. In this study, we use wireless network logs and computing resources login logs to construct a more precise survey of ownership and usage of devices on campus. We propose a methodology using Machine Learning to extract usage patterns, including using a mixed-effect regression model to analyze the impact of Class Standing (i.e., freshman/first-year, sophomore, junior, senior) and the student's academic division to explain usage patterns. Overall, we report on ownership and usage profiles for personal laptops, phones, tablets, loaned laptops, and computer lab desktop computers. The results invite consideration of student device provision strategies that address a University's desire for robust analytics data while mitigating disproportionate impacts to students.

Keywords: Student Success, BYOD, Mobile Computing, Student Technology Use, Mobile Learning, Student Engagement and Interaction, Teaching with Technology

1 APPLYING DEVICE USAGE DATA TO TECHNOLOGY DECISION-MAKING

In this study, we used the directory information and University's wireless networks and computing resources logs of 19,675 degree-seeking undergraduate students during Fall term 2017 (OSU Institutional Research Office Enrollment Report). Our research evaluates the observable ownership of devices and student device usage behavior on campus to help understand the potential impact of a transition to a BYOD model. Our findings show that detectable ownership, that is, the actual rates of student device ownership and usage while onsite, are substantially lower than those estimated by the Pew Research Center (Smith, et al., 2011; 88% of undergraduates reported owning a laptop and 59% owning a desktop) or ECAR (Dahlstrom et al., 2016; laptop ownership >93%). We also explored the correlation between device ownership and usage of campus computing resources. Whereas the BYOD model reflects a belief that most students already own devices and do not use computer labs regularly. Our report corrects this, in the context of Oregon State University, by showing that a non-negligible proportion does not observably use a laptop while on campus

Our study characterized the ownership landscape among undergraduates for personal laptops, phones, and tablets to reveal the on-campus usage patterns for each type of device and the correlation between personal laptop usage on campus or ownership and computer lab usage.

1.1 Observable laptop ownership

Our analysis revealed that 64.75% connected a personal laptop computer; we also found a churn of a few thousand students who did not "bring back" their devices. For each student, we counted how many days per week they used a device and then calculated the proportion who used that device one day vs. two days, etc., per week. The red line represents the cumulative number of unique users observed. We exploited these vectors of counts to mine behavior patterns for laptops and phones by using the k-means clustering algorithm to generate profiles of weekly usage of these devices, including what percentage of students used a laptop on campus for various numbers of days per week, what proportion of students never used a laptop on campus, and what proportion were identified as using neither a personal laptop nor a university provided computer. Figure 1 reflects our analysis of laptop ownership; due to space constraints, phone ownership data will be included in the poster.

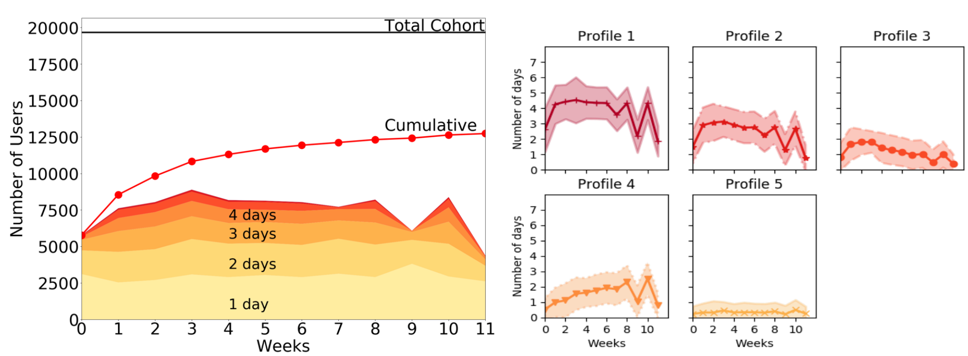


Figure 1: Number of Unique Students per Number of Days They Used a Personal Laptop on Campus' Wireless Network; Wireless Network Usage Profiles - Fall 2017

We identified five weekly profiles that shape our understanding of students' behaviors. They indicate that while most students own a laptop, few (approximately 20%) bring it to campus more than 2 days per week, on average.

REFERENCES

Dahlstrom, E., et al. (2016, October). *ECAR Study of Undergraduate Students and Information Technology, 2016*. Research report. Louisville, CO: ECAR. Retrieved from https://library.educause.edu/~/media/files/library/2014/10/ers1406-pdf.pdf?la=en

Smith, A., Rainie, L., & Zickhur, K. (2011). *College students and technology.* Pew Research Center. Retrieved from http://www.pewinternet.org/2011/07/19/college-students-and-technology/

Knowledge Building Discourse Analyzer: A New Tool to Explore if Students Build their Knowledge Through Discourse

Ahmad Khanlari
University of Toronto
a.khanlari@mail.utoronto.ca

Abstract: Knowledge creating dialogue is central to Knowledge Building, because learners construct their knowledge, express their opinions, values and feelings through discourse. Therefore, it is very important to realize if the discourse is moving toward a knowledge objective. The purpose of this study is to develop a tool in order to assess knowledge building discourses and examine if individuals' knowledge is improving over time or not.

1 INTRODUCTION

The aim of Computer Supported Collaborative Learning (CSCL) is to embrace technology to facilitate collaborative learning. The general belief about CSCL is that constructing individual's knowledge is primarily achieved through discourse (Stahl, 2003). Knowledge Building, which is described as the production and continual improvement of ideas of value to a community (Scardamalia & Bereiter, 2003), holds an even stronger belief in the role of discourse in learning: "the state of public knowledge in a community only exists in the discourse of that community, and the progress of knowledge just is the progress of knowledge-building discourse" (Scardamalia & Bereiter, 2006, p. 12). In Knowledge Building communities, students are not only engaged in advancing their own knowledge but in advancing the state of knowledge in their community. The purpose of this study is to develop a tool in order to analyze the discourse progress and realize if students' knowledge is improving over time as a result of engaging in Knowledge Building discourse.

2 METHOD AND PLAN OF DATA ANALYSES

In order to assess individuals progress, I examine alignment between individual's discourse and the curriculum as students progress through the semester. It is expected that the more students learn in a field, the similarity between their talk and experts talk is higher (Lave & Wenger, 1991). Therefore, alignment between the curriculum and students' discourses can be considered as an indicator of scientificness, which is a result of sociocultural learning. The participants of this study include Grade 1 students who used a web based discourse medium, called Knowledge Forum®, to advance their understanding of water and water cycle through social interactions. In order to find the similarity index (i.e. alignment) between individuals knowledge and the curriculum, a matrix which includes words frequencies is created. To create such a matrix, Python's Natural Language Processing Tool Kit (NLTK) is used, and employing Part of Speech (PoS) technique, nouns and names in students' notes and the curriculum were extracted, as it is assumed only nouns and names may represent concepts (Kopainsky, Pirnay-Dummer, & Alessi, 2012). The rows in this matrix are the terms and their frequencies in curriculum and individuals' notes (Table 1). Then, using the Cosine Similarity metric (Leydesdorff, 2005) the similarity index between individuals' notes and the curriculum is computed. This analysis has been done for each week in order to track individuals progress over time. The similarity index ranges from zero (indicating decorrelation) to one (meaning exactly the same). Values between zero and one indicate intermediate similarity or dissimilarity.

Table 1: A sample of the word-frequency matrix

Document/Term	cloud(s)	droplets	water	planets	atmosphere	altostratus
Curriculum plan	25	6	9	1	1	1
Student 1	8	1	13	0	0	0
Student 2	7	0	4	0	0	0

3 PRELIMINARY RESULTS, CONCLUSION, AND FUTURE DIRECTIONS

The results of running the tool on three students' discourses is shown in Figure 1. As Figure 1 shows, the similarity index between students' discourses and the curriculum is increasing over time. For example, while the similarity index between S3 discourses and the curriculum was only 0.27 in the first week, it increased to 0.43 and 0.54 in Week 2 and Week 3, respectively. The increasing similarity index can be an indicator of idea improvement over time, which is required in Knowledge Building communities.

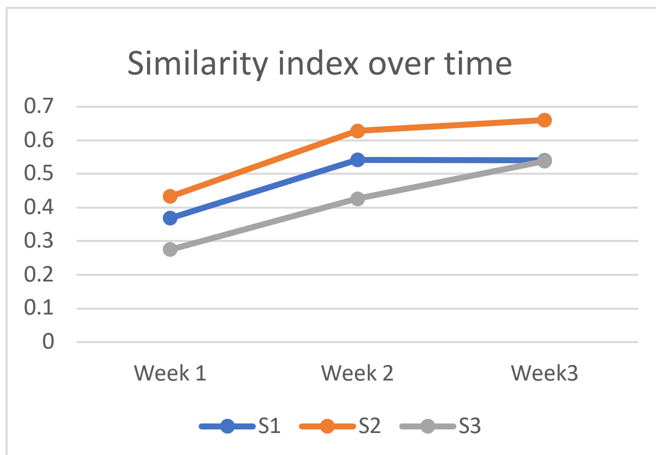


Figure 1: Similarity index for 3 randomly-selected students

This tool not only can provide teachers with an overview of students' progresses, but also can be embedded in Knowledge Forum dashboard in order to provide students with formative feedback about their progresses over time. The future direction of the study includes running the tool over a rich data set, as well as running the tool over the community knowledge (in contrast to individuals' knowledge) to examine if the community knowledge is improving over time or not.

REFERENCES

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

Leydesdorff, L. (2005). Similarity Measures, Author Cocitation Analysis, and Information Theory. *JASIST, 56*(7), 769-772.

Scardamalia, M. & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97-115). Cambridge University Press.

Scardamalia, M., & Bereiter, C. (2003). Knowledge Building. In *Encyclopedia of education* (1370-1373). New York: Macmillan Reference, USA.

Stahl, G. (2003). Building collaborative knowing: Elements of a social theory of learning. In J.-W. Strijbos, P. Kirschner & R. Martens (Eds.), *What we know about CSCL in higher education*. Amsterdam, NL: Kluwer.

In-class Programming Exercise Support System with Real-Time Error Information to Learners

Keiko Shigeta
Graduate School of Media and Governance, Keio University
kiki@sfc.wide.ad.jp

Achmad Husni Thamrin
Graduate School of Media and Governance,
Keio University
husni@sfc.wide.ad.jp

Jun Murai
Faculty of Environmental Information,
Keio University
jun@wide.ad.jp

ABSTRACT: This paper proposes a learning support system for elementary programming classes by displaying the coding errors to the learners when they execute their codes. We conducted experiments at an elementary JavaScript programming course at our campus and found that students who had basic understanding of the syntax and understood the displayed errors performed well in the class. We plan to extend our system for a better learning analytics.

Keywords: Real-time Automatic Feedback; Support for Instructors; Programming in Group Education; In-Class Support System; Elementary Programming Course;

1 INTRODUCTION

This research focuses on leading learners in programming classes such as in universities to fix their coding mistakes by themselves with little guidance from the instructors to minimize the discrepancies between parallel classes by reducing the dependency to the teaching skills of the instructors in different classes. We proposed a system, which is in principle an online IDE, and evaluated it based on the collected data.

2 PROPOSAL AND RESULTS

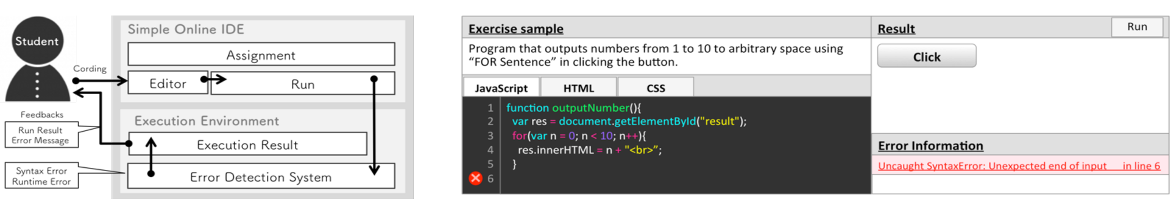


Figure 1: System architecture and its UI. The system displays both the syntax and runtime errors.

This research proposes a system for programming classes using automatic feedback of coding errors to the learners to increase their comprehension level within the lecture time. Figure 1 shows the system architecture and its user interface. When a learner executes his/her code, the system displays and records the results and the errors. The system prominently displays the errors to make it easier for both the learners and the instructors to identify the programming mistakes, hence the instructors can quickly give the proper advises to as many learners as possible.

We evaluated the system at a compulsory elementary JavaScript programming class at our campus with 65 students. The students took 2 tests (before and after) and did 4 programming exercises in the class. Based on the collected data, among others, we found how the students struggled through the exercises as the exercise difficulty increases (Fig. 2a). For example, there were 6 students who gave up after 2 tries in a basic exercise (Q1), which we may interpret as a lack of interest or motivation. While for a difficult exercise (Q3), they were doing more tries before giving up, and the data for Q4 is due to the class time was over.

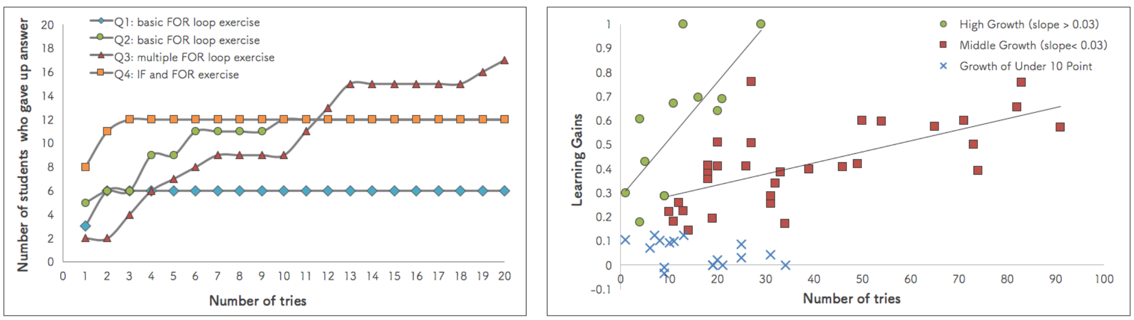


Figure 2: (a) The cumulative number of students who give up after certain number of tries. (b) We classified the students into high, medium and high growth based on their learning gains vs the number of tries.

We classified the students based on their learning gains (Fig. 2b) in relation with the number of tries. The high growth students are the ones who had high learning gain with a small number of tries. We looked into the error messages in order to understand the differences between these classes and we found that the low growth students did not really understand the JavaScript syntax, while the medium and high growth students mainly made mistakes in function calls or algorithm. As the medium and high growth students had similar error types, we believe that the high growth students had better understanding of the displayed error messages.

3 DISCUSSION

This system only displayed the error messages as feedback to the learners in real-time and tracked few types of learners data in a programming class, but it did not provide any insight to the class instructors in real-time. As this work is only at the preliminary stage, we are planning to extend our system for better learning analytics as well as testing its efficacy in supporting programming courses in our campus.

REFERENCES

Youngs, E. A. (1974). Human errors in programming. *International Journal of Man-Machine Studies*, *6*(3), 361-376.

Blikstein, P. (2011, February). Using learning analytics to assess students' behavior in open-ended programming tasks. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 110-116). ACM.

Fu, X., Shimada, A., Ogata, H., Taniguchi, Y., & Suehiro, D. (2017, March). Real-time learning analytics for C programming language courses. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 280-288). ACM.

Demonstrations

AWA-Tutor: A Platform to Ground Automated Writing Feedback in Robust Learning Design

Antonette Shibani
University of Technology Sydney, Australia
antonette.shibani@gmail.com

Increasingly, the importance of aligning learning analytics with learning design is being understood as a way to uphold its core aim of improving educational practices, while also collecting meaningful data about learner's activities that can be interpreted in context (Lockyer, Heathcote, & Dawson, 2013). In light of this, a writing analytics tool "AWA-Tutor" has been developed that integrates analytics with pedagogy. AWA-Tutor is a web-based tool developed as an extension of the Academic Writing Analytics (AWA) tool that provides automated feedback on students' writing based on rhetorical structures in the text (Shibani, Knight, Buckingham Shum, & Ryan, 2017).

AWA-Tutor extends AWA, by scaffolding an entire writing improvement activity. Students are guided through a series of tasks, such as understanding the instructor's rubric, improving a sample text, reviewing exemplar improvements, self-assessing their work, and reflecting on the quality of the automated feedback. The tool is designed in a modular fashion to support the learning design of an instructor, who can select the task components to be included, and personalize the feedback experience for different students. AWA-Tutor captures detailed activity traces: the time taken by students to complete certain tasks of the activity, snapshots of drafts at customizable time intervals, students' requests for automated feedback and the feedback received, and feedback survey responses. Thus, the process of drafting and revising, which were previously hard to study, are now reconstructable for subsequent analysis in other tools such as R, for which a suite of analyses have been developed.

AWA-Tutor has been evaluated with undergraduate law students in authentic classroom settings, tackling tasks co-designed with the Law academic, performing the activity individually or in pairs. Both student and instructor feedback have been positive regarding the usage of this tool over the two semesters the intervention was run, although there is certainly scope for improvement.

Demonstration movie: https://www.youtube.com/watch?v=K212XabCL5w&feature=youtu.be

Keywords: AWA-Tutor, writing analytics, tool, learning design, pedagogy integration

REFERENCES

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439-1459.

Shibani, A., Knight, S., Buckingham Shum, S., & Ryan, P. (2017). *Design and Implementation of a Pedagogic Intervention Using Writing Analytics.* Paper presented at the 25th International Conference on Computers in Education, New Zealand.

Scaling MOOC Discourse Analysis with In Situ Coding

Hui Liang, Brennan Sieck, Christopher Brooks
University of Michigan School of Information
lianghui@umich.edu, bsieck@umich.edu, brooksch@umich.edu

ABSTRACT

Educational discourse research is a broad and interdisciplinary field. The methods used in the field are very diverse as well, which include analysis of video in which learners are engaging with one another, social network analysis, and text discourse analysis. The last of these methods, text discourse analysis, is commonly used by learning analytics and learning science researchers to study the cognitive behaviors of MOOC learners and peer interaction in online learning environments. To perform such analysis, researchers typically hire coders who would go into online discussion fora and annotate student writing based on the coding scheme selected in the research study.

Through discourse analysis, researchers aim to create a more personalized learning experience for learners by providing real-time support and feedback. To solve “the information overload and chaos” in MOOC discussion forums, Wise et al. (2016) developed a bags-of-word model to separate content-related threads from non-content-related ones. In an exploratory study, Kovanović et al. (2016) built a tree-based model to automate the classification of forum messages into the four different levels of cognitive presence defined by the Community of Inquiry framework with the idea that such models can help provide pedagogical guidance from the discussion fora alone.

However, discourse data in MOOCs remains largely unexplored for several reasons. We point to three issues which we aim to address here in particular, (a) the lack of technical knowledge in getting MOOC discourse data, (b) the logistic issue researchers have in getting access to MOOC discussion fora, and (c) the separation of the coding activity from the context of the discourse, as most researchers load data into third-party applications, such as Excel or NVivo, for coders to annotate manually.

In this demo, we present *Innotate*, a Chrome extension specifically designed for all researchers regardless of their technical background. We specifically built *Innotate* as a Chrome Extension because it is easy to install and requires zero configuration on the client side. With this tool, researchers can easily update coding scheme of their choice and collect discourse data at a large scale across multiple MOOC courses. This enables researchers to evaluate the generalizability of their results more efficiently. Coders can view and annotate messages within its context where the messages are being created by learners. Once the coders highlight some text in a forum message, a pop-up box will appear and ask the coder to select an appropriate code for the highlighted text. The coder can then add the annotation and click the submit button in the user interface of the extension.

To access the demo video, please visit https://youtu.be/Op4PIB8k828 .

ProTuS - Interactive Learning Analytics Tool

Boban Vesin, Katerina Mangaroska, Michail Giannakos, Hallvard Trætteberg
Norwegian University of Science and Technology, Trondheim, Norway
{vesin, mangaroska, michailg, hal}@ntnu.no

ABSTRACT: ProTuS is personalized and interactive e-learning tool that provides different visualization reports to support learning process. It offers several, mostly programming courses and it is currently used and evaluated at the Norwegian University of Science and Technology for 1st year's Web technologies course.

Keywords: e-learning, programming courses, visualized reports

1 PROGRAMMING TUTORING SYSTEM

ProTuS[1] is the programming tutoring system developed at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. It provides learners with personalized courses from various domain (Vesin, Ivanović, Klašnja-Milićević, & Budimac, 2011). Currently, ProTuS offers different presentation methods and adapts user interface using techniques of tag-based recommendation (Vesin, Ivanović, Budimac, & Pribela, 2008).

To support concepts of open learner model and learning analytics, several visualization reports are offered or planned for development in ProTuS. These reports present information about test results, learners' progress, personal and group activities, fulfillment of learning objectives, etc. Additional visualizations are planned to support learning analytics integration and aggregation of learning-related data across multiple sources.

ProTuS has been used and tested by 10 teaching assistants and 66 first year bachelor students from the *Web technologies* course at NTNU. The experiments showed that the system provided useful insights for further implementation of learning analytics component of ProTuS.

REFERENCES

Vesin, B., Ivanović, M., Budimac, Z., & Pribela, I. (2008). Mile - Multifunctional integrated learning environment. In *MCCSIS'08 - IADIS Multi Conference on Computer Science and Information Systems; Proceedings of e-Learning 2008* (Vol. 2).

Vesin, B., Ivanović, M., Klašnja-Milićević, A., & Budimac, Z. (2011). Rule-based reasoning for building learner model in programming tutoring system. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7048 LNCS, pp. 154–163). https://doi.org/10.1007/978-3-642-25813-8_17

[1] ProTuS home page: https://protus.idi.ntnu.no/, Demonstration video: https://youtu.be/RQ-KoLFlwLg

Learning Analytics in Schools

Learning Analytics in Schools

Simon Buckingham Shum
University of Technology Sydney, AUS
Simon.BuckinghamShum@uts.edu.au

Abelardo Pardo, Peter Reimann, Kathryn Bartimote-Aufflick
The University of Sydney, AUS
firstname.lastname@sydney.edu.au

ABSTRACT: The data and analytics revolutions are disrupting and already transforming many sectors in society: finance, health, shopping, politics. Data is not new to education, but for many, it is still challenging to articulate the connection between the potential of using data to support decision making, and the day-to-day operations occurring in learning environments. At this inaugural LAK workshop for the K-12 sector, school leaders, teachers, data analysts, academics, policy makers and all other interested parties were invited to join a professional learning and development day focused on the practical applications of Learning Analytics in school (K-12) education. Drawing on national and international expertise, speakers include innovative school leaders and teachers, school data analysts, university researchers, software and consulting companies.

Keywords: K-12, primary school, secondary school, high school, data, analytics, feedback, teacher dashboards, student dashboards, 21st Century Competencies

This workshop (https://latte-analytics.sydney.edu.au/school) was developed for newcomers, as well as for those who already knew a about Learning Analytics, or were already using it in the classroom.

The objective was for participants to leave with a deeper understanding of:

- the diverse forms that Learning Analytics can take, and especially how technology extends this far beyond conventional school data to create better feedback;
- how such data is being used by school leaders to support strategic reflection;
- how new kinds of data are being used by teachers to support their practice;
- the practicalities of initiating such work in one's own school.

We are indebted to our speakers for their contributions, invited to provide a snapshot of the diversity of work now underway in schools, from diverse perspectives. We encourage you to browse their briefings and explore their work more deeply.

This schools-focused event is a new initiative for the international LAK conference. It is hoped that colleagues will forge new professional connections, catalyzing new work to advance the responsible, effective use of analytics to advance teaching and learning.

Program

8:55	**Welcome and Initial Remarks** Susi Steigler-Peters (CEO of Research Australia Development & Innovation Institute)
9:00 Ignite Talks Data for School Leaders	**What is "Learning Analytics" and why a Schools Day?** Simon Buckingham Shum (University of Technology Sydney) **Learning Analytics: For Insight into a School Environment** Gary Molloy (St Aloysius College, Sydney) **Architecting Whole School Data for Whole Pupil Development** Ruth Deakin Crick (University of Technology Sydney)
10:00	Refreshments
10:30 Ignite Talks Data for Teachers and Students	**Tracking and Visualising Student Effort: A Practical Analytics Tool for Student Engagement** Robin Nagy (Educational Consultant, Sydney, Australia) **Teachers Co-designing Innovations and Sharing Data to Improve Outcomes** Jojo Manai (Carnegie Foundation for the Advancement of Teaching, USA **edQuire: Providing teachers with insights into students' ICT competencies** Michael Cejnar (edQuire, Australia) **Formative learning analytics to foster 21st century competencies in Singapore secondary schools** Jennifer Tan & Elizabeth Koh (National Institute of Education, Singapore)
11:30 Roundtables	**Roundtable discussions** The speakers and additional invited practitioners will be available for Q&A about their work. Delegates are invited to browse the briefings circulated in advance, and are free to move between groups in 'unconference' mode. The discussions will be divided into two topics: • **Data for School Strategy and Leadership Decisions.** The use of data for school leaders to improve processes, gain insights to inform strategic decisions, etc. • **Data to Improve Teaching and Learning.** The use of data for day-to-day teaching duties. Personalising content, feedback, etc.
12:30	Lunch
13:30 Workshop	**Getting real. What conditions make learning analytics work?** Small group discussion to envision how these techniques can be effectively deployed in schools.
14:30 Closing Panel	**Closing Panel: From Vision to Reality** Closing reflections on the day from the experts and open discussion Facilitator: Susi Steigler-Peters
15:15	*Close, Refreshments, Informal Conversations*

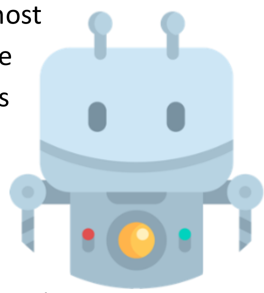
Self-Directed Student Learning: A Digital Learning Infrastructure for Self-Directed Learning

Ruth Crick
University of Technology Sydney
Ruth.Crick@uts.edu.au

ABSTRACT: Complex problem solving, critical thinking and creativity are the three most important capabilities for thriving in the Fourth Industrial Revolution. [i] These are not traditionally developed through legacy learning and development systems (human or digital) because they require real-world, purposeful contexts, the ability to work across silos, new measurement models and courageous leadership. Learning design for teachers is about creating the conditions where students can take responsibility for their own learning by invoking their own passion and purpose and the agency to pursue these through a learning journey in contexts where the outcome is not known in advance. This session will focus on the challenges and opportunities of building such a Digital Learning Infrastructure and will use live examples from the new Learning Emergence Learning Journey Platform which is in production in release 1.0 with a group of schools in the UK. In April this year a group of schools in the Hunter NSW will be using it sponsored by Hunter Water Corporation who are using the same Learning Journey platform as a vehicle for cultural transformation as they move into the uncertainty and challenges of infrastructure resilience and sustainability for the future of the region.

Keywords: self-directed learning; complex systems; digital infrastructure; measurement models

1 Why does this matter?



Complex problem solving, critical thinking and creativity are the three most important capabilities for thriving in the Fourth Industrial Revolution.[i] These are not traditionally developed through legacy learning and development systems (human or digital) because they require real-world, purposeful contexts, the ability to work across silos, new measurement models and courageous leadership. Learning design for teachers is about creating the conditions where students can take responsibility for their own learning by invoking their own passion and purpose and the agency to pursue these through a learning journey in contexts where the outcome is not known in advance.

2 What are we measuring?

The most important unit of change is the story and identity of the learner – not the teacher, the curriculum or the measurement model. Legacy systems tend to privilege the content of the curriculum, a reductionist measurement model and the teacher as agent of change. The challenge for learning analytics is to build a digital infrastructure based on a data architecture which provides a 'single view of the learner', where data belongs to the learner and can be used, one student at a time, in real-time, for better decision-making as they navigate their way through complex problems to solutions that matter to them. This is sometimes described as a call to move towards Education 3.0 – a challenging worldview shift from a top down, individualist and dualistic worldview (Education 1.0) towards an integral, participatory and wholistic one. For a discussion about these ideas see the first Handbook for Learning Analytics and a chapter called **Layers, Loops and Processes.** [ii]

3 What is this session about?

This session will focus on the challenges and opportunities of building such a Digital Learning Infrastructure and will use live examples from the new Learning Emergence Learning Journey Platform which is in production in release 1.0 with a group of schools in the UK. In April this year a group of schools in the Hunter NSW will be using it sponsored by Hunter Water Corporation who are using the same Learning Journey platform as a vehicle for cultural transformation as they move into the uncertainty and challenges of infrastructure resilience and sustainability for the future of the region.

4 The *Learning Journey* platform

The purpose of the Learning Journey Platform is to enhance self-directed learning capabilities, and thus the resilient agency, of students, teachers and leaders and schools across the world. It provides scaffolding support for people in authentic enquiry learning journeys which contribute measurably to data-informed local solutions that matter *and* empower self-directed, resilient learners. 'Learning power' is a term which describes this approach. [iii] Resilient people are a pre-requisite for resilient and sustainable practices at all levels of society. See this link for an introduction to the Learning Journey Platform.[1]

4.1 Loops – feedback and feedforward

Rapid feedback of meaningful data is key to enhancing self-directed learning. The Learning Journey Platform hosts the *CLARA* learning power assessment tool, the *TESA* teacher development tool for pedagogy which supports deep student engagement and Angela Duckworth's *Grit* survey. Feedback to the user is immediate and provides a framework for reflection – 'backwards' towards identity and purpose and 'forwards' to a particular purposeful outcome.

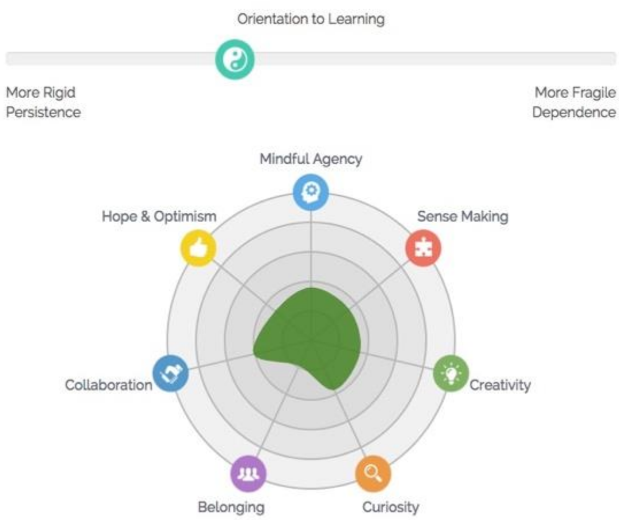


Figure 1: CLARA – 8 Dimensions of Learning power

[1] https://www.youtube.com/watch?v=FJ_LHp6sC3Y

The Learning Journey Platform aggregates anonymised data in real time for teachers and leaders to interrogate in different ways. This capability is possible because of the underlying data architecture which allows for a 'single view of the learner'. The data belongs to the learner and they can take their learning journeys with them from school to school and on to University and into the work place.

Figure 2: TESA: Teaching Expertise, Scaffold & Analytics

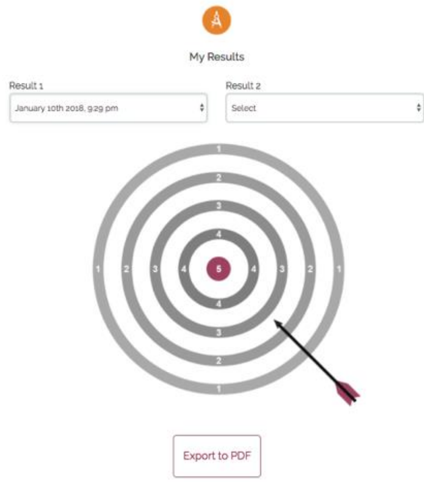


Figure 3: GRIT – Passion and Perseverance

4.2 Processes – the learning journey

A key design principle underpinning the learning journey platform is that learning is a journey that begins with a purpose and moves towards an outcome or 'performance' of some sort. When a student defines and owns their own purpose – the *why* - they are at the beginning of resilient agency. They need to use their learning dispositions – their learning power – to understand themselves as learners and to figure out *how* to move towards their purpose. The *what* is the data, information, experience and new knowledge they need to identify, collect, curate and re-construct in order to achieve their purpose. This is a familiar enquiry cycle for most educators – the key difference here is the emphasis on purpose and agency and self-directed navigation.

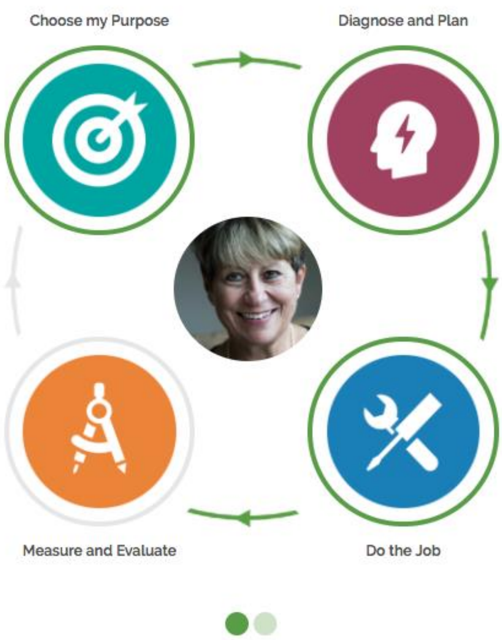


Figure 4: The Learning Journey design

The learning journey metaphor is simple and yet profound in terms of mind-set shifts. A person leads a journey, you can be on your own or with others, there's a terrain, a map if you're lucky, challenges, diversions and a destination. Journeys have endings and beginnings and way points, and come in all shapes and sizes.

The Learning Journey Platform builds on best practice in data architecture from financial services in customer journeys and uses AI to support the individual learner in navigating their learning. Whereas in the commercial world the focus is on the 'next best action' in the world of learning the focus is on the 'next best question'. Dialogue and discourse are at the heart of learning.

4.3 Layers – students, teachers, leaders, system leaders

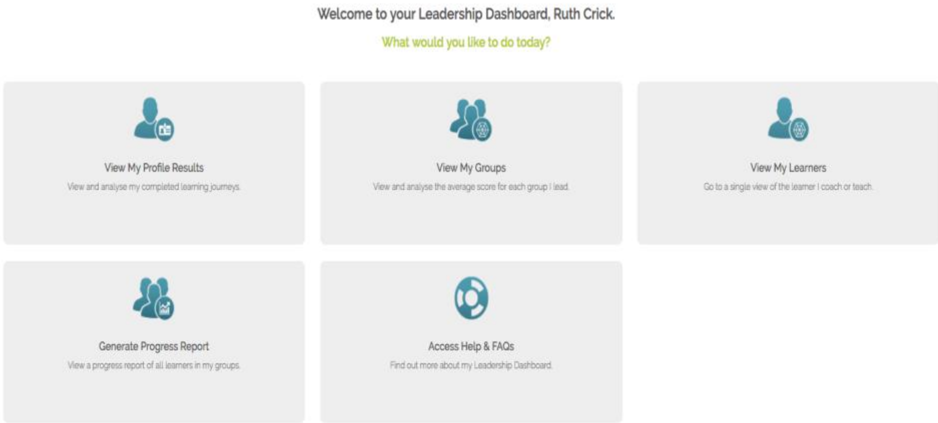


Figure 5: Leadership Dashboard

Schools are complex living systems which are multi-layered. We know how important teacher professional learning is – you can't give what you haven't got. Moving towards education 3.0 means to be part of a worldview shift which is happening around us because of the challenges of life in the 21C.

A worldview shift of this type is uncomfortable and challenging. It's best encountered and managed through deep professional learning – for leaders and teachers. The Learning Journey Platform captures the data, analyses it and returns aggregated anonymised data as feedback to teachers and leaders for more focused interventions and better decision making. Personal data is only viewed by another person with explicit permission.

5 What next?

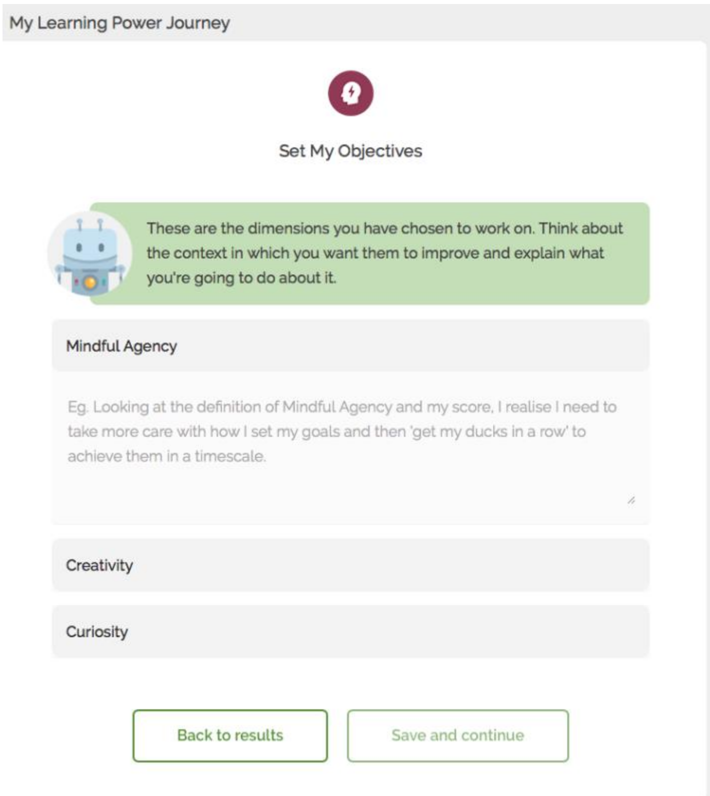


Figure 6: Buddy asking questions

The focus for the next stage of the Learning Journey Platform is on enhancing the use of AI to support purposeful conversations – enhancing, not replacing, the face to face relationships of trust, affirmation and challenge that are at the heart of learning. Buddy already asks questions and 'calls time' for reflection at key junctures in each journey and he'll get cleverer as time goes by. The second focus is on developing support and scaffolding for a whole authentic enquiry project.

The Learning Journey Platform is available for use by schools and HE in this phase of development. Its capability to collect and integrate data around rapid cycles of enquiry make it an ideal candidate to support professional learning and improvement science approaches to educational transformation. Its partnership with Declara – social learning and knowledge curation - mean that through the INSIGHTS tab capability users can access 'knowledge pathways' – units of relevant learning material which sit within Declara. The potential for scaling up professional learning across geographies and time is significant.

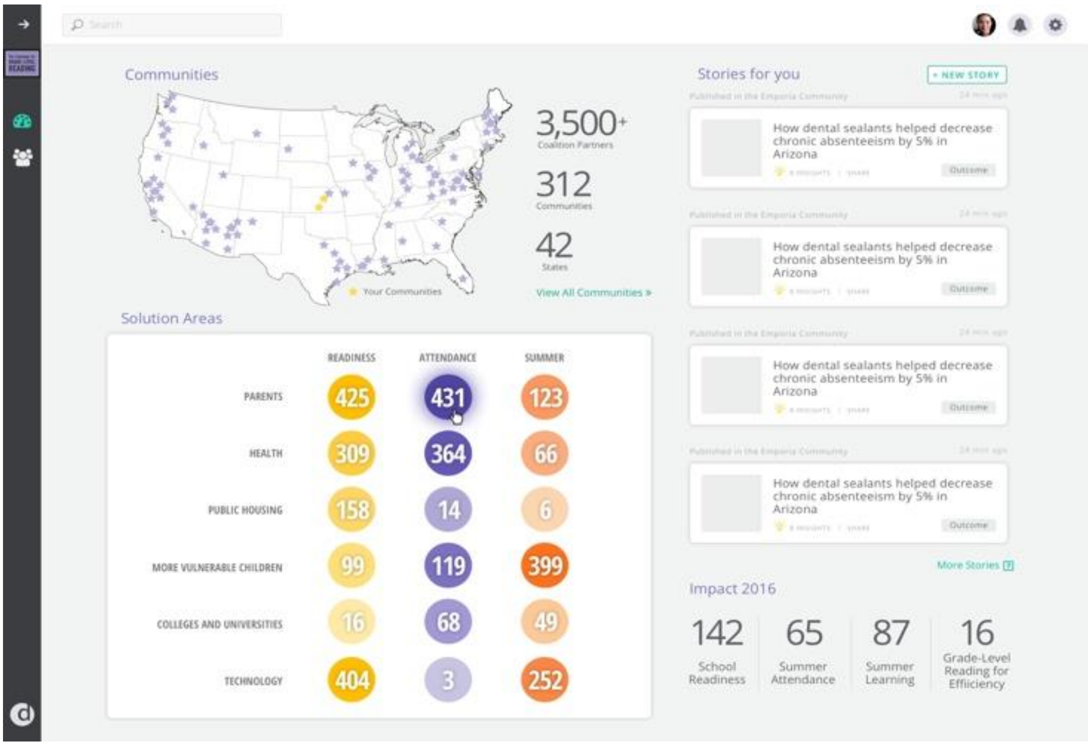


Figure 7: Declara – social learning and knowledge curation at scale (http://declara.com)

6 Business models

This sort of education innovation requires new business models that allow for collaboration, innovation and evolution. The Learning Emergence Partnership is developing a wholistic approach where the same learning design principles are used in industry for cultural transformation both in terms of employees and customers. In between education and industry there is 'community engagement' and 'vocational education'. Our vision is to make this work accessible for all schools, working with both industry and philanthropy. Learning Emergence has an asset locked Foundation to ensure this.

References

i https://www.weforum.org/agenda/2017/09/skills-children-need-work-future-play-lego

ii Crick, R. 2017, 'Learning Analytics: Layers, Loops and Processes in a Virtual Learning Infrastructure' , in G. Siemens & C. Lang (eds), *Handbook of Learning Analytics,* 1st Edn, Society for Learning Analytics Research, pp. 291-307. https://solaresearch.org/hla-17

iii Deakin Crick, R., Huang, S., Ahmed Shafi, A. & Goldspink, C. 2015, 'Developing Resilient Agency in Learning: The Internal Structure of Learning Power', *British Journal of Educational Studies*, vol. 63, no. 2, pp. 121-60. http://dx.doi.org/10.1080/00071005.2015.1006574

Tracking and Visualising Student Effort: A Practical Analytics Tool for Student Engagement

Robin Nagy
Educational Consultant, Sydney, Australia
Robin.Nagy@EffortTracking.com

Abstract: To what extent do good academic results predict success beyond high-school? What other traits should we be measuring and nurturing at school? In the world beyond the classroom, dispositional characteristics such as grit, resilience and self-discipline are increasingly becoming valued by employers and by research, as more reliable indicators of success than academic results alone. But how do you measure and nurture these dispositional character traits? This briefing showcases a school-wide Student Effort Tracking project which has been implemented in two Sydney high-schools over 8 years and has successfully helped to improve student motivation for learning in all cohorts, creating high-quality data-driven coaching conversations between students, teachers and parents.

Keywords: Effort, growth mindset, grit, resilience, dispositional learning analytics, visualization, 21st century skills, intrinsic motivation, student tracking.

1 OVERVIEW – THE CASE FOR MEASURING AND TRACKING EFFORT

The rationale behind a school-wide focus on 'Effort', rather than solely on academic achievement, is to improve intrinsic motivation for learning in all students, by explicitly identifying and recognising the behavioural and learning dispositions which promote growth mindsets and lead to academic development and improvement. One of the ways in which this data is reported to staff and students is using a dynamic bubble-chart to display student progress over time: https://vimeo.com/168306314.

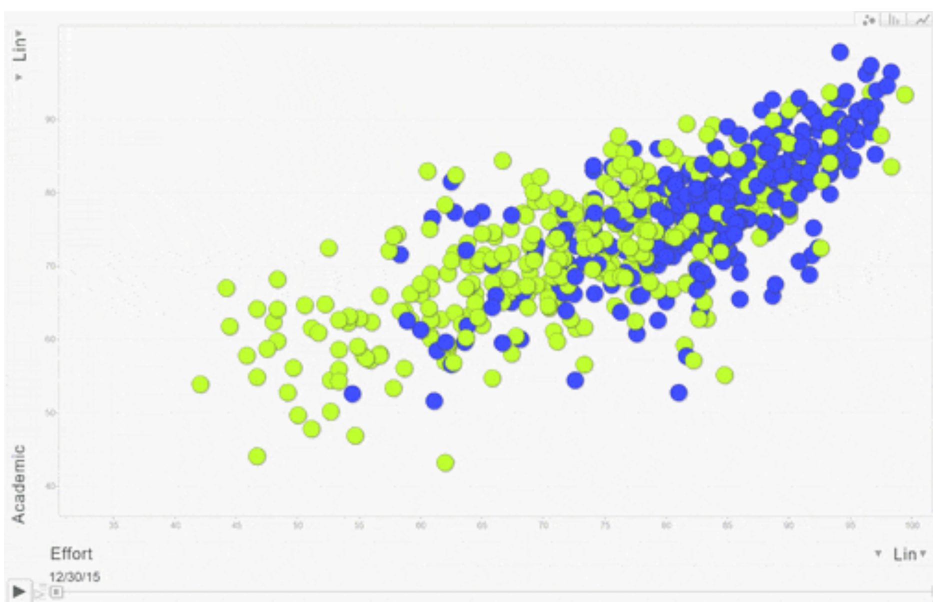


Figure 1: A Still Frame from the Dynamic Bubble-Chart showing Academic Achievement against Effort with boys in green and girls in blue.

Fundamental to ensuring buy-in from all stakeholders, has been agreement of a shared set of expectations and standards for 'Effort-grading', including the creation and development of agreed criteria and rubrics. This has been achieved through regular staff and student dialogue and evaluation and refinement of the salient behavioural and learning dispositions which comprise 'Student Effort'.

Effort Tracking Rubric					
	5- Outstanding	4- Very Good	3- Good	2- Fair	1- Unsatisfactory
Behaviour	Classroom Conduct and Attitude, Politeness and Respect, Consideration for the Learning of Others				
	Proactively models positive classroom behaviour and attitude at all times, avoids distraction and shows respect and consideration for others. Is polite and courteous at all times.	Consistently demonstrates good behaviour and attitude conducive to learning and avoids distractions in class.	Usually demonstrates a positive attitude in class and is rarely distracted.	Generally shows a positive attitude in class but is sometimes distracted or inconsiderate of the learning of others.	Rarely exhibits conduct and attitude appropriate for a conducive learning environment.
Diligence	Self-discipline, Self-reflection, Independent Motivation, Persistence, Conscientious Application to Classwork and Homework				
	Demonstrates an excellent approach to all activities in class and at home, presenting work to the best of his/her ability at all times and bringing all required equipment to class. Is independently motivated and disciplined and takes pride in the quality of all work produced, frequently exceeding expectations of conscientiousness and persistence.	Completes all work to a high personal standard in a timely manner and fulfils all expectations for coursework. Brings all equipment to class. Demonstrates a self-disciplined approach to all activities and often independently persists when academically challenged.	Usually completes work to a good personal standard, brings equipment to class and demonstrates self-discipline in application to coursework.	Shows some self-discipline in completing most coursework with a reasonable level of application.	Rarely fulfils expectations with regard to self-discipline, conscientiousness and application to coursework.
Engagement	Classroom Focus, Communication (Verbal and Body Language), Personal Presentation and Punctuality, Participation and Contribution in Groups and Class				
	Consistently demonstrates the highest standards of attention and focus in class, contributing where appropriate to group or classroom forums and/or demonstrating active listening skills at all times. Is always punctual and well-presented.	Actively listens to all teacher explanations and instructions and where appropriate, participates in group and class forums. Is punctual and well-presented.	Usually demonstrates good focus in class, listening to teacher instructions and explanations and appropriately participating in group and class forums. Is usually punctual and well-presented.	Is generally well-focused and on-task in class, participating from time to time in group class forums.	Is rarely focused in class and often off-task.

Figure 2: ***An example of the Effort Tracking Rubric***

An overall 'Effort Score' is then created by averaging all individual subject effort-grades and scaling to yield a number from 20 to 100. This 'Effort Score' is then tracked against the student's academic achievement from term to term, and presented in the dynamic bubble chart, set (anonymously) against the background of all other students. Importantly, students also grade their own effort using the same approach (without seeing their teachers' grading) and their 'Effort Score' can then be compared to that of their teachers.

Following publication of these effort grades, at the start of each term, teachers have targeted student-led coaching conversations with all students based on their effort scores from the previous term, and students use the 'bubble-chart' and quantitative subject-specific information to set goals for the term ahead within the context of 'Effort'.

Although this rubric continues to be discussed and refined, it is often in the analysis of teacher-student effort-assessment discrepancy that shared expectations are re-aligned. Indeed, there is evidence to suggest that the precise syntax of the rubric has little overall effect on the distribution of grades and that both teachers and students adopt a 'global impression' approach to their assessment. Nevertheless, the ongoing development of the rubric provides an important process for articulating a set of shared standards and expectations surrounding the learning environment.

1.1 Overview of the theoretical and methodological approach

The theory and development of this project are set out extensively in a peer-reviewed paper (Nagy 2016a) published in the Journal of Learning Analytics Special Section on Learning Analytics for 21st Century Competencies in September 2016. A summary of this work can be viewed in my presentation at UTS (Nagy, 2016b).

Schools seek to maximise the best possible academic outcomes for their students, but these are usually determined systemically, through assessment via high-stakes summative testing. However, a results-driven success-focus can paradoxically lead to a decline in achievement for some students and a widening of the gap between higher and lower achieving students; this is due to the detrimental effect on student wellbeing and intrinsic motivation for learning (McDonald, 2001; Harlen & Deakin Crick, 2003). Students who lack innate 'academic buoyancy' (Martin, 2010) see their lack of success in academic assessment as evidence to support a fixed mindset; that diligence has no effect on 'smartness' and they lose confidence in their own capacity to learn (Black & Wiliam, 1998; Harlen & Deakin Crick, 2003).

1.2 Outcomes of the Effort Tracking Project

There is considerable evidence to suggest that the process is making a positive difference to student motivation for learning with a significant improvement in average Effort grades in every cohort over a two year period. Moreover, on several occasions, a dramatic decline in a particular student's effort score has highlighted (sometimes previously unknown) pastoral issues and allowed timely interventions to occur. Comparison of student and teacher grading can often reveal students with perfectionist tendencies as well as those who lack an objective sense of their own character. In these cases, appropriate interventions and targeted coaching conversations can be tailored to suit individual students' needs.

There have been interesting indications that further analysis of boys' and girls' effort grading may yield useful information which could further inform teachers' practices. For instance, in all cohorts, the discrepancy between (teacher assessed) boys' and girls' effort scores is significantly greater than the discrepancy between their respective academic achievement scores. This would suggest that perhaps teachers' classroom expectations may be naturally biased towards female behavioural and learning dispositions. Alternatively, it may suggest that girls are on average more likely to identify and adhere to teacher expectations and/or appear more visible in doing so. Reflecting on whether or not our shared expectations of classroom behaviour, diligence and engagement are equally beneficial for both boys' and girls' learning is one important consequence of this analysis.

By focusing on the processes rather than the outcomes of learning, the school's 'success-focus' can be intentionally shifted towards these more nurturing and developmental dispositions under the umbrella term 'Effort'. In this way, all students can see a more immediate indication of 'success' by monitoring progress in their 'Effort Grades'. This promotes the adoption of a growth mindset when viewed in the context of the dynamic bubble chart, where students have visual reinforcement of the positive, but delayed, correlation between effort and academic achievement. In addition, student intrinsic motivation is fostered with a positive effect not only on academic achievement and student wellbeing, but on lifelong-learning traits and character development.

2 CONCLUSIONS AND RECOMMENDATIONS

Research indicates that increasing testing does not raise academic standards and promotes an extrinsic rather than intrinsic motivation for learning (Harlen & Deakin Crick, 2003). Rather than nurturing a joy of lifelong learning, this 'results-driven focus' emphasises distinct ability-divisions which promotes 'fixed-mindsets' in students, teachers and parents. The result is to create an academic climate where failure is seen as a reinforcement of inability, rather than a challenge to be overcome, and one in which students' anxiety levels increase, often with a detrimental effect on their performance and wellbeing.

By comparing students with each other based on their effort, rather than their achievement alone, we subtly shift the systemic 'success-focus' onto qualities which promote a growth-mindset in all students and develop important 'non-cognitive' character traits such as persistence and resilience. The engagement and continuing professional development of teachers is critical to embedding and sustaining a project of this sort.

This project shows that, although challenging, evaluating and quantifying student effort is possible, and that it is in the dynamic tracking processes and conversations surrounding this formative form of assessment, where many of the main benefits are to be found, rather than in the momentary snapshots and finite, summative effort scores themselves.

REFERENCES

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74. http://dx.doi.org/10.1080/0969595980050102

Bryk, A. (2015). *Learning to improve: How America's schools can get better at getting better*. Boston: Harvard Education Press.

Buckingham Shum, S., & Deakin Crick, R. (2012). Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics**.** *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (LAK'12), 92–101. http://dx.doi.org/10.1145/2330601.2330629

Costa, A. L., & Kallick, B. (2009). *Habits of mind across the curriculum: Practical and creative strategies for teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.

Covey, S. R. (1989). *The 7 Habits of Highly Effective People: Powerful Lessons in Personal Change*. New York: Free Press.

CLARA (2016). Crick Learning for Resilient Agency: Learning Emergence LLC. [Web site]. Retrieved from http://clara.learningemergence.com

Deakin Crick, R., Huang, S., Ahmed-Shafi, A., & Goldspink, C. (2015). Developing resilient agency in learning: The internal structure of learning power. *British Journal of Educational Studies, 63*(2), 121–160. http://dx.doi.org/10.1080/00071005.2015.1006574

Duckworth, K., Fielding, G., & Shaughnessy, J. (1986). *The relationship of high school teachers' class testing practices to pupils' feelings of efficacy and efforts to study*. Portland, OR: University of Oregon.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Jackson, R. C. (2014). Concentration. In R.C. Eklund, G. Tennenbaum (Eds.), *Encyclopedia of sport and exercise psychology* (Vol.1), (pp. 159—161). Thousand Oaks, CA: Sage Publications.

Harlen, W., & Deakin Crick, R. (2002). A systematic review of the impact of summative assessment and tests on students' motivation for learning. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education, 10*(2), 169–207. http://dx.di.org/10.1080/0969594032000121270

Hobsons. (2014). *Student Feedback and Progressive Reporting* (Report). Melbourne: Hobsons Australia Pty Ltd. Retrieved from http://www.bps.sa.edu.au/__files/f/16997/Hobsons student feedback and progressive reporting June14.pdf

Kellaghan, T., Madaus, G., & Raczek, A. (1996). *The use of external examinations to improve student motivation.* Washington, DC: AERA.

Martin, A. J. (2010). *Building classroom success: Eliminating academic fear and failure.* London: Continuum.

McDonald, A. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology, 21*, 89–101. http://dx.doi.org/10.1080/01443410020019867

Nagy, R. P., (2016a). Tracking and Visualising Student Effort: Evolution of a Practical Analytics Tool for Staff and Student Engagement. *Journal of Learning Analytics,* 3(2), 164-192. http://dx.doi.org/10.18608/jla.2016.32.8

Nagy, R. P., (2016b). The Development of a Student Effort Tracking Model. Seminar, Connected Intelligence Centre, University of Technology Sydney (June 8th 2016). https://utscic.edu.au/events/niccic-redlands-school-8-june-2016

Park, D., Tsukayama, E., Goodwin, G. P., Patrick, S., & Duckworth, A. L. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal and intellectual competencies in children. *Contemporary Educational Psychology,* 48, 16-27.

Perkins-Gough, D. (2013, September). Educational leadership: The significance of grit — A conversation with Angela Lee Duckworth. *Educational Leadership*, 71(1), 14–20. http://www.ascd.org/publications/educational-leadership/sept13/vol71/num01/The-Significance-of-Grit@-A-Conversation-with-Angela-Lee-Duckworth.aspx

Pollard, A., Triggs, P., Broadfoot, P., McNess, E., & Osborn, M. (2000). *What pupils say: Changing policy and practice in primary education*. London: Continuum.

Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low achieving pupils to high stakes testing. *Educational Evaluation and Policy Analysis*, *23*, 197–228.

Schunk, D. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*, 207–231. http://dx.doi.org/10.1080/00461520.1991.9653133

Smith, L. H., & Kays, T. M. (2010). *Sports psychology for dummies "cheat sheet."* Mississauga, ON: John Wiley & Sons Canada.

Teachers Co-designing Innovations and Sharing Data to Improve Outcomes

Jojo Manai
The Carnegie Foundation for the Advancement of Teaching
manai@carnegiefoundation.org

ABSTRACT: A chasm has been growing for some time between our rising aspirations and what our schools are able to routinely achieve. And this chasm is greatest for our most disadvantaged students and institutions. So this has now formed as one of the great social justice issues for our time. Carnegie propose Networked Improvement Communities (NICs) as structure of colleagueship that combines research, professional, improvement knowledge in creative ways to develop effective changes for improvement. In this presentation we will use The Carnegie Math Pathways as an example of how educators and researchers work together to improve outcome using the principles of improvement science.

Keywords: Networked Improvement Community, Behavior Modeling, Statistical Analysis, Growth Mindset, Just-in- time Interventions, Productive Persistence

1 OVERVIEW OF LAST DECADE'S EFFORTS IN CLOSING CHASM

1.1 Performance Management strategies

A strength of this method is that it has focused our attention on data. Today we have much greater transparency around student outcomes including disparities in these outcomes among various groups. This is a plausible strategy if you thought this was principally a problem of effort or that educators were just not focused on the right topics. On the other hand, if it was a problem of educator learning, of not knowing how actually to do better, the performance management is lacking in a critical regard. There is no detailed working "theory of practice improvement" in this strategy.

1.2 Evidence-based Practice Movement

This approach has brought enhanced theoretical discipline and greater analytic rigor to the work of improvement. The overall corpus of work here, however, is quite modest. At best, what these studies tell us is that some interventions can work because they must have worked somewhere for somebody for the positive effect to emerge. But they do not tell us whether they will actually work for you, and under your circumstances.

1.3 School-Based Learning Communities

In a sense, this strategy stakes out a polar opposite position to the evidence-based practice movement. Field trials tend to focus on evaluating commercial products, say a new curriculum, a specific pedagogic practice or a technology. The evidence derived, here again, is in essence, an average result, and the influence of local context in all of this tends to fade into the background. In contrast, communities of practice take local context very seriously as they focus in on the day to day problem of improving work in specific classrooms, schools, and districts.

2 INTRODUCING THE IMPROVEMENT PRINCIPLES

The six principle of improvement are:

1. Be problem-focused and user-centered
2. Attend to variability
3. See the system
4. Embrace Measurement
5. Learn through disciplined inquiry
6. Organize in networks as practical, scientific communities

Improvement science starts with investigating the specific problems we need to solve (principle 1). This means focusing in on the unsatisfactory variability in outcomes that we observe (principle 2) and seeing how our educational systems (principle 3) create these outcomes, often unintentionally. At the core of improvement research are rapid iterative cycles of testing possible change ideas against data, revising, retesting and refining (principles 4 and 5). And then to tackle the larger, more complex and persistent problems we confront, we join together in improvement networks (principle 6). While our individual capacities may be modest, working together we can achieve much more (Figure1).

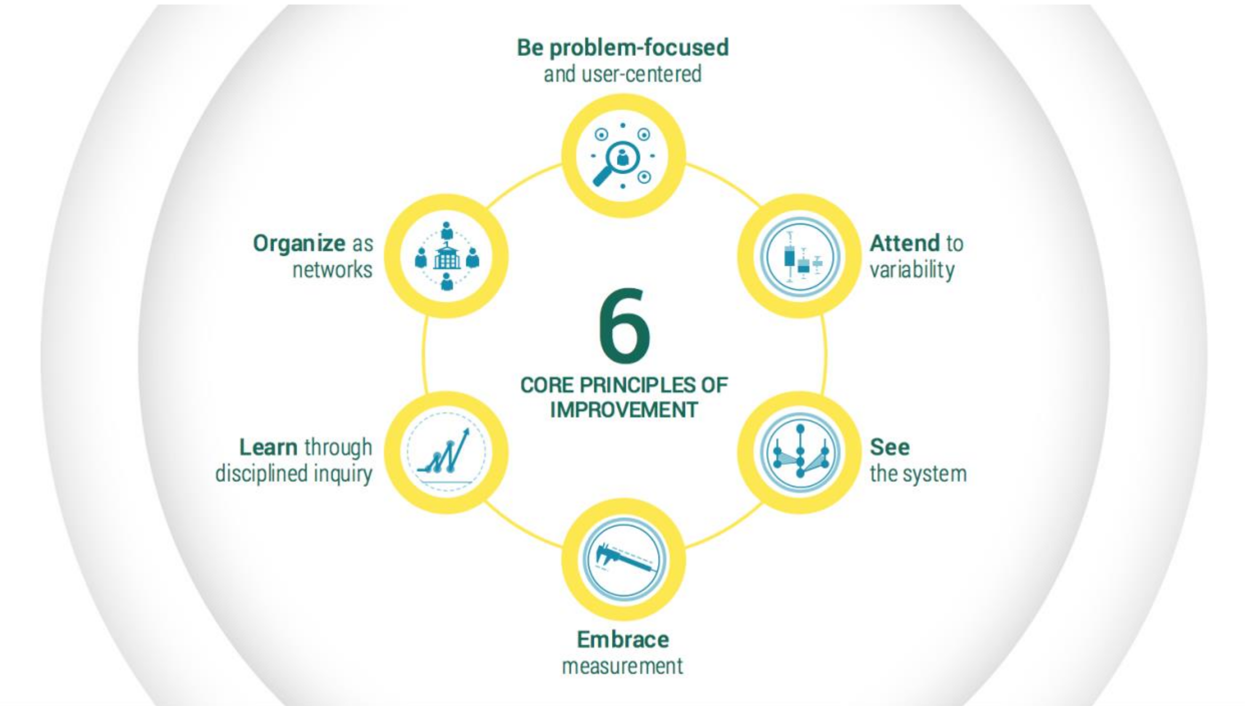


Figure 1: The Six Improvement Principles

3 THE CARNEGIE MATH PATHWAY AS AN EXEMPLAR OF A NIC

Approximately 60 percent of the nation's incoming community college students are referred to at least one developmental math course, of which 80 percent will not earn college-level math credit within three years [1]. Without achieving college math credit, they cannot transfer into four-year degree programs or qualify for entry into preparation programs in a wide range of occupational-technical specialties. As a result, millions of students each year fail to acquire essential mathematics skills and are unable to progress toward their career and life goals.

To address this national problem, the Carnegie Math Pathways (CMP) program was developed and implemented through Networked Improvement Communities (NICs) involving college faculty, administrators, researchers, designers, and content experts [2, 3, 4]. Statway is one of the CMP initiatives designed as a year-long course that allows students to simultaneously complete their developmental mathematics and college-level statistics requirements to receive college credit. A causal analytic study using a propensity score matching technique [7] with hierarchical linear modeling (HLM) approach [5, 6] confirmed that Statway tripled the success in half the time across two different cohorts, and that the effect held across the gender and race/ethnicity groups.

4 PRODUCTIVE PERSISTENCE AS A DRIVER TO IMPLEMENT CHANGE

4.1 Definition

- Students have skills, habits, and know-how to succeed in college
- Students believe they are capable of learning math
- Students feel socially connected to peers, faculty, and the course
- Faculty actively support productive student mindsets and engagement.
- Stronger student (and faculty) engagement yields higher students results.

4.2 Classroom Measurements and Interventions

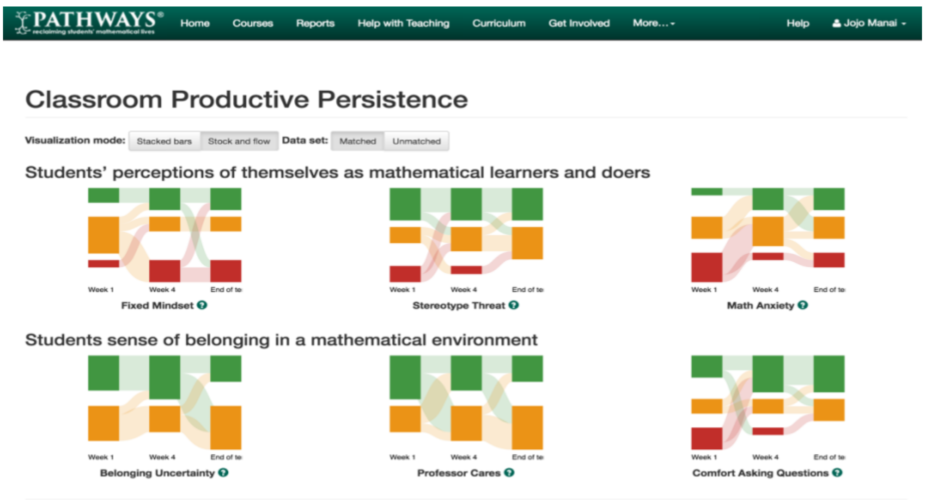


Figure 2: Classroom Productive Persistence Report

REFERENCES

[1] Bailey, T., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. Economics of Education Review, 29, 255–270.

[2] Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. T. Hallinan (Ed.), Frontiers in sociology of education (pp. 127-162). New York, NY: Springer.

[3] Bryk, A. S., Gomez, L. M., Grunow, A., LeMahieu, P. G. (2015). Learning to improve: How America's schools can get better at getting better. Cambridge, MA: Harvard EducationPress.

[4] Dolle, J. R., Gomez, L. M., Russell, J. L., Bryk, A. S. 2013). "More than a network: building communities for educational improvement." In B. J. Fishman, W. R. Penuel, A. R. Allen, and B. H. Cheng (Eds.), Design-based implementation research: Theories, methods, and exemplars. National Society for the Study of Education Yearbook. New York, NY: Teachers College Record.

[5] Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. Educational Evaluation and Policy Analysis, 27, 205-224.

[6] Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage Publications.

[7] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70, 41-55.

Real Time Learning Analytics of Computer Use in K-12 Classrooms

Michael Cejnar, ChienHui Tancy Kao
EdQuire Project, FIC Technology Pty Ltd, Sydney
m.cejnar@edquire.com.au

ABSTRACT: Australian K-12 classrooms have adopted 1:1 computer use, however, academic results have been inconsistent arguably due to a lack of ICT skills and distraction. We describe a novel real-time learning analytics tool giving teachers and educators feedback on actual computer usage in order to improve teaching and policy. We used an app transmitting student activities to a cloud-based AI algorithm to educationally analyse in real-time, students' classroom use of computers. Data from 549 year 7-12 students from 4 schools were collected over 9 months. Results: Computers were used for 20 minutes per lesson; overall 17% was Off-task, but for 20% of students, a full 30% was Off-task. Boys spent marginally more time Off-task, mostly on gaming, while girls were most distracted by streaming videos. Distractibility Index (DI) combining Off-task with task switching was steady in Years 7-10, but dramatically improved in Years 11-12. On-task use remains mostly for word processing and content delivery. Search-engine usage analysis shows 99.5% use of Google and with only 3% of searches using any advanced search tools. Distractibility was reduced by a significant 31% in a subgroup of students given feedback of their DI data. Our findings demonstrate the feasibility, utility and we believe necessity for objective measurement and learning analysis of 1:1 computer use in classrooms to develop strategies and means for ensuring their effective use.

Keywords: Student engagement, Digital Learning, Learning Analytics, AI, Computer behaviour

1 INTRODUCTION

K-12 classrooms in Australia are adopting 1:1 computer programs to facilitate learning and instil ICT skills (Crook & Sharma, 2013). Results are however mixed and inequitable (NAP-ICT 2014) with increasing computer use arguably associated with distraction and lower educational outcomes, related to a Second Digital Divide separating skilful versus distracted ICT learners (Cameron, Bennet & Agostinho, 2011). Little objective data is available on how computers are actually used in class, and teachers lack the ability to detect and monitor their students' engagement. We describe here our first results from a novel classroom computer learning analytics tool, EdQuire, intended to make computer learning visible to teachers and educators in order to improve computer-aided teaching and policy.

Using a background computer agent transmitting student activities to a cloud-based AI-based learning analytics algorithm, we categorized and analysed the educational relevance of student computer usage data in real time and gave teachers a 'glanceable' colour coded display showing their student's engagement. Here we demonstrate the utility of this analysis by describing extent of computer usage in class, the proportions of On-task versus Off-task use, nature of activities, a brief overview of search engine use by students and the effect on distractibility of giving behavioural feedback to students.

2 METHODS

2.1 School & Students

Classroom computer usage data were collected from four schools with 1:1 computers between Feb and Oct 2017. The study included 549 students (286 boys and 263 girls) from year 7 to 12. Data from a total of 3,961 student lessons with computer use were analysed from 4 schools. Computer usage data from a total of 21,454 student lessons were analysed and described.

2.2 Data collection

An EdQuire software agent was installed on student computers, which securely transmitted student computer usage data in real time to our secure cloud service. Data connections were established with school time tabling and student information systems. Data was collected, with student and parent knowledge, only during lesson times and stored anonymously with student identifying data held separately for security. Activities collected included applications, application's window title, visited websites and browser tab titles. Keystroke frequency and use of clipboard was recorded. No typed content was recorded except for search engine text contained in URLs. Duration spent on each website or application lasting more than half a second was recorded. Activities without user input for 2 minutes were reclassified as Idle.

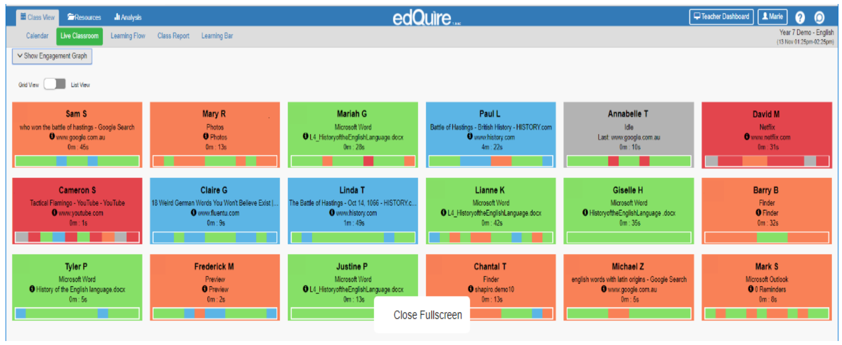


Figure 1: Teacher EdQuire classroom console showing teacher-assigned resource On-task (green), student-discovered On-task (blue), entertainment Off-task (red), unclassified/ambiguous task (orange), and Idle (grey). Horizontal striped bars show the history.

2.3 Data real-time display

Teachers were given an internet console for use in lessons, displaying in a colour code the students' current 'on-taskness' and its history, allowing them to see at a glance classroom engagement and identify outlying, struggling or unchallenged students who potentially needed support (Figure 1).

2.4 General analysis method

Categorizing on-taskness: Student computer usage activities were first categorized in real time by a machine learning algorithm as educational or non-educational (accuracy 93%). Activities were then further automatically contextualized, as teacher-assigned On-task (green) if concordant with a list of schoolwide or teacher-entered lesson resources, or other student-discovered On-task educational activities (blue). Ambiguous activities (orange) were resolved by educational experts. For this analysis, blue and orange tasks were grouped into and all referred to as On-task (green).

Distractibility Index (DI): Off-task time and frequent task-switching contribute to distraction (Kraushaar and Novak 2010). We thus defined a Distractibility Index as a ratio of time On-task, versus time Off-task weighted by the number of activity switches between the two. This index ranges from 0 (no distraction) to 1 (completely distracted).

Student Feedback: a subgroup study of 153 students were given access to their on-taskness and DI data for 8 weeks and their behaviour compared to prior 8 weeks.

3 RESULT

3.1 Students' active computer use in lessons and On-task vs Off-task

Computers were used in 84% of lessons for at least 2 minutes. The average computer use time in these lessons was 20 minutes per lesson, occupying 45% of a lesson's 40 to 50 minutes. Typical usage spanned 10 to 35 minutes with 5 outlier students apparently spending entire lessons actively using their computer (Figure 2-A). Overall, students spent 17% of their class computer time Off-task but for 20% of students, a full 30% was Off-task (Figure 2-B).

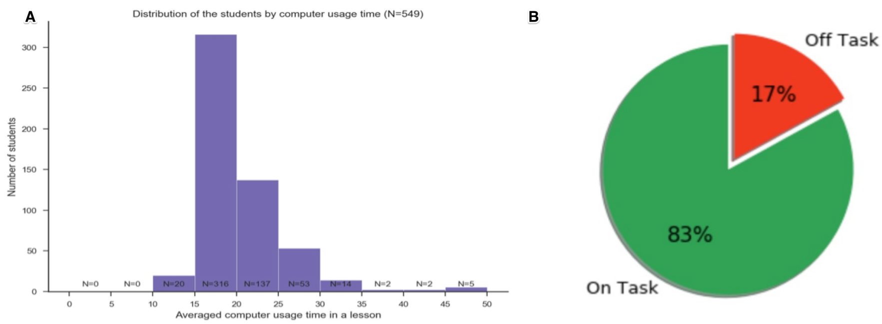


Figure 2: A) Distribution of students by computer usage time in a lesson and B) Proportion of On-task vs. Off-task duration

3.2 Computer usage behaviour for boys and girls

Computer use in lessons was similar for both genders (20.6 and 20.2 minutes for boys (N=286) and girls (N=263), respectively). Boys had statistically significant but only marginally higher Off-task time of 19% per lesson compared to 15% for girls.

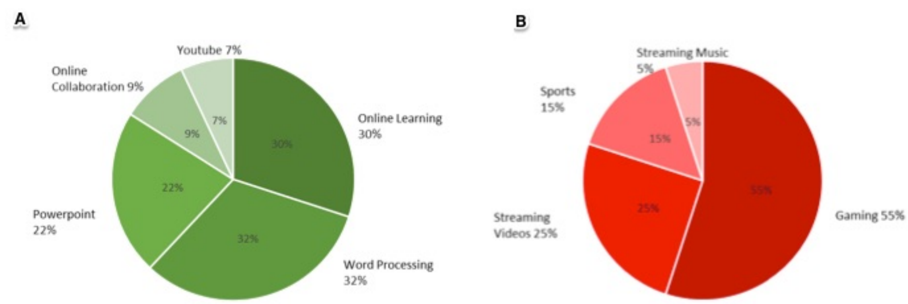


Figure 3: A) Proportion of top 10 On-task activities, and B) proportion of top 10 Off-task activities for all students.

3.3 Nature of On-task and Off-task activities

We calculated the top 10 activities by duration, including apps and websites, in the On-task and Off-task categories. Overall, students spent most of their On-task time on word processing (i.e. MS Word), online learning (Learning management system (LMS), online learning websites), PowerPoint, collaboration (i.e. Google docs), and email (Figure 3A). Students spent most of their

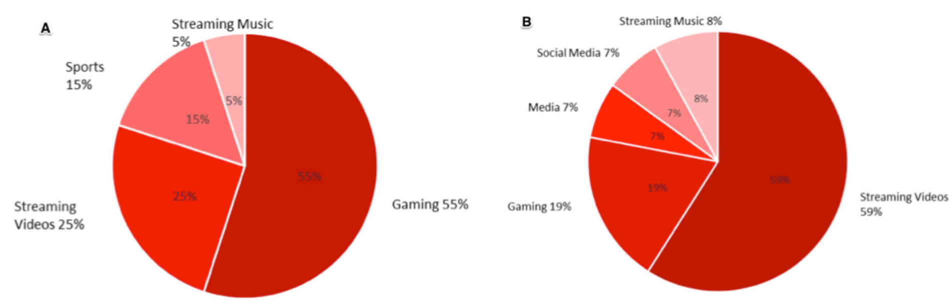


Figure 4: A) Proportion of top 10 Off-task activities for boys, and B) proportion of top 10 Off-task activities for girls.

Off-task time on computer gaming (i.e. Flappy Bird and the pseudo-educational Coolmath-games), streaming videos (via Netflix, YouTube, Movie Player), sport (i.e. NBA, FIFA sites), and streaming music (Figure 3B). In a subset of data with attributable gender, boys and girls had similar On-task behaviour pattern to all students. In their Off-task time, boys visited mostly games sites, sport and then streaming videos and music (Figure 4A) while girls visited mostly streaming videos and less games sites (Figure 4B).

3.4 Distractibility by grade Year

We calculated the average Distractibility Index for each student per lesson, ranging from 0 to 1 (1=completely distracted). Distractibility was similar across Years 7 to 10 with a dramatic drop in distractibility in Years 11 and 12 (Figure 5).

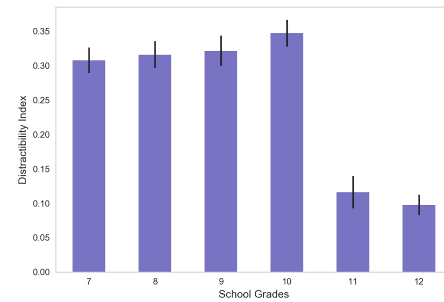


Figure 5: Distractibility in Years 7 to 12.

3.5 Internet search skill assessment

We investigated the Internet searching behaviour of students. We analysed which search engines students used, and then their use of available search tools such as the Boolean AND, OR, NOT, quotation marks and others (such as +, -, ~, :, or, and, site) to refine their search. We found that Google accounted for 99.5% (Table 1) and less than 3% of searches utilised any advanced search tools.

Table 1: Proportion of Internet search engine usage

Internet search engine	Proportion of usage
Google	99.5%
Bing	0.43%
Yahoo	0.04%
wolframalpha	0.01%

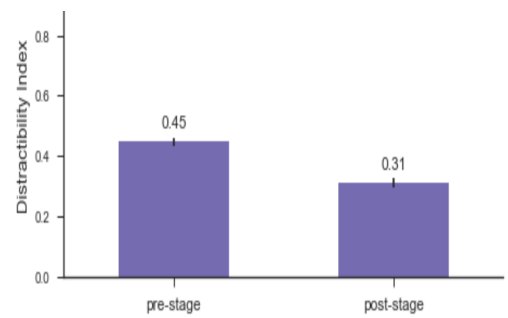


Figure 6: Proportion of regular and advanced search frequency.

3.6 Student Feedback

Students in this sub-study looked at their feedback an average of 3 times a lesson, consuming 19 seconds and reduced their DI by 31% from 0.45 to 0.31 (Figure 6), which persisted across the 8 weeks.

4 DISCUSSION AND CONCLUSION

While the overall 17% of the 20 minutes of computer time spent Off-task is modest, 20% of students were potentially problematic, spending a 1/3 of their computer time Off-task. A quick look at their data would reveal to the teacher if they were finished and bored or perhaps struggling and tailor the right intervention. The overall good results may due to the highly motivated students, awareness of computer monitoring during our trial, or of course, the benefit of EdQuire real-time class display used by teachers confirmed by logs, but we are unaware of other similar objective data sets for comparison.

The preponderance of word editing and on-line LMS activities suggests a persistent use of computers for delivery of content and as writing tools as found by Crook & Sharma, 2013. Presentation software accounting for 22% of time suggests time spent on creating or just reading presentations; analysis of our keyboard and clipboard activity data will help us distinguish this. The large fall in distractibility in Years 11 and 12 may be due better self-regulation with age or due to students focusing on their HSC; inspection of their On-task and Off-task activities will likely give more clues.

Internet searching was ubiquitous and almost all using Google engine. The minimal use of search engine tools may be a reflection of lack of skills or a preference for iterative use of searches to create a trail converging on the answer, which may actually be useful by adding context (White & Huang 2010). Further analysis of search trails and quality of landing page will further assess digital literacy.

The student interest in using feedback and the consequent persistent reduction in distractibility is our most exciting result and supports the well documented utility of timely formative feedback for self-regulation and learning. It is the subject of our next report.

In conclusion, we believe this is the first study in the body of K-12 literature, which analyses actual computer use in classrooms and makes it visible in real time to teachers and educators. We have that shown such a Learning Analytics tool can provide measurement and understanding of computer use and misuse in classrooms in order to explain where and why it fails to deliver educational outcomes. This in turn enables the formulation and validation of strategies and practices for effective computer use, such as informing teacher interventions or giving usage feedback to students. Critically, such a measurement tool is then needed to monitor the effective application of such strategies to individuals and cohorts of students, in order to ensure we teach all our students the ICT skills necessary to provide equitable outcomes. Our study demonstrates that K-12 real-time learning analytics is feasible, practical, and scalable and that further development and application of platforms like EdQuire will provide the 21st Century ICT teaching and assessment tools needed by our teachers, educationalists and policy makers as well as by parents and students themselves.

REFERENCES

Cameron, T., Bennett, S. & Agostinho, S. (2011). ICT literacy and the second digital divide: Understanding students' experiences with technology. In T. Bastiaens & M. Ebner (Eds.), Proceedings of EdMedia: World Conference on Educational Media and Technology 2011 (pp. 3392-3397). Waynesville, NC: Association for the Advancement of Computing in Education (AACE).

Crook, S., & Sharma, M. (2013). Blooming Heck! The Activities of Australian Science Teachers and Students Two Years into a 1:1 Laptop Program Across 14 High Schools. *International Journal of Innovation in Science and Mathematics Education, 21(1),* 54-69.

Kraushaar, J., & Novak, D. (2010). Examining the Effects of Student Multitasking With Laptops During the Lecture*. Journal of Information Systems Education, 21(2),* 241-251.

NAP-ICT (2014). Fraillon, Julian; Schulz, Wolfram; Gebhardt, Eveline; and Ainley, John, (2015). "*National Assessment Program : ICT Literacy Years 6 & 10 Report 2014" Sydney: Australian Curriculum, ACARA* http://research.acer.edu.au/ict_literacy/11

White, R., & Huang, J. (2010). Assessing the scenic route: Measuring the value of search trails in web logs. Proceedings of SIGIR, January 1, 19-23.

Formative Learning Analytics to Foster 21st Century Competencies in Singapore Secondary Schools

Jennifer Pei-Ling Tan and Elizabeth Koh
National Institute of Education, Nanyang Technological University, Singapore
jen.tan@nie.edu.sg, elizabeth.koh@nie.edu.sg

ABSTRACT: Recent research on the assessment of 21st century competencies (21CC) has accentuated the need to move beyond traditional modes of assessment towards contemporary approaches that can better capture and reflect the protean nature of 21CC and its constituent skills and dispositions, and the interdependent and dynamic interactions that occur in the process of learning. Towards this end, WiREAD and My Groupwork Buddy, have been designed to leverage learning analytics (LA) to provide rapid formative assessment of students' 21CC development as both process and outcome. These have been trialled in multiple secondary schools in Singapore and provide ways of assessing 21CC as a formative endeavour to productively inform and scaffold everyday learning and teaching. Benefits, challenges and pathways forward are briefly mentioned.

Keywords: Learning analytics, 21st century competences, critical literacy, teamwork, dashboards, formative assessment

1. LEARNING ANALYTICS AND 21ST CENTURY COMPETENCIES

It is commonly posited that assessment drives learning. In Singapore, the education fraternity recognizes the importance of augmenting current assessment practices to better foster 21st century competencies (21CC) beyond academic achievement. Recent research on the assessment of 21CC has accentuated the need to move beyond traditional modes of assessment towards contemporary approaches that can better capture and reflect (i) the protean nature of 21CC and its constituent skills and dispositions, and (ii) the interdependent and dynamic interactions that occur in the process of learning (Tan, Choo, Kang, & Liem, 2017). So how can we make visible or 'see' 21CC in the classroom? What are some ways that 21CC can be measured and characterized as they occur naturalistically in students' peer interactions during acts of learning? These imperative questions continue to confound educators today.

Our work engages directly with these questions by examining ways of assessing 21CC as a formative endeavour that can productively inform and scaffold everyday learning and teaching. Jennifer and Elizabeth are lead investigators of WiREAD and My Groupwork Buddy (MGB) respectively, two projects that have been designed to leverage learning analytics (LA) to provide rapid formative assessment of students' 21CC development as both process and outcome.

1.1 WiREAD

WiREAD aims to enhance students' 21st Century Skills in English Literacy through a web-based collaborative reading and learning analytics environment (Figure 1).

Figure 1: WiREAD's learning environment

Two chief learning designs are its multimodal social dialogic learning and dynamic visual learning analytics (Figure 2). *Wi*READ allows teachers to (i) choose/integrate appropriate multimodal textual resources, (ii) embed critical reading pedagogical scaffolds based on Multiliteracies pedagogy, Paul-Elder's 'wheel of critical reasoning' (Paul & Elder, 2001), as well as dialogic indicators of collective creativity and criticality (Tan, Caleon, Jonathan, & Koh, 2014), and (iii) continuously monitor learning progress and adapt pedagogical strategies to stimulate students' deep multimodal engagement and rich peer-to-peer critical interactions around texts, both during and beyond formal English language class time.

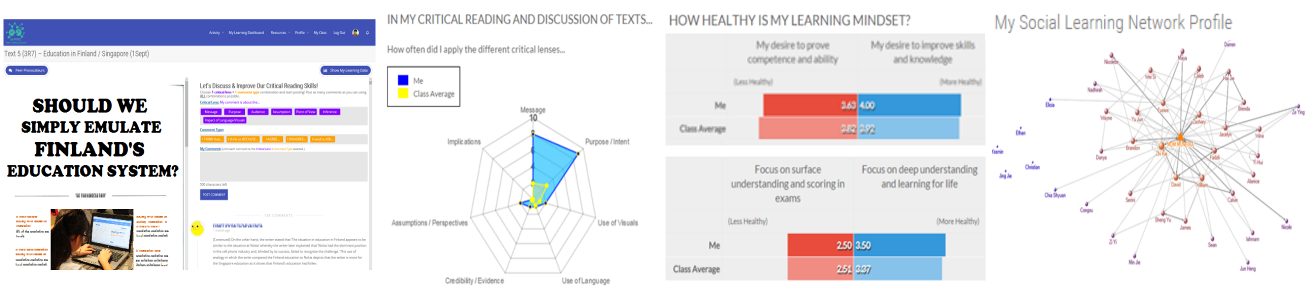


Figure 2. WiREAD's multimodal social dialogic learning and dynamic visual learning analytics

Twenty-two texts and scaffolding prompts (Critical Reading Lens and Collaborative Talk types) have been co-developed as ICT-based lesson packages through a tight partnership between the research team, one seed innovation school, and policy officers of the Education Technology Division at the Ministry of Education. To date, approximately 1000 secondary students across multiple schools in Singapore have used WiREAD to enhance collaborative critical reading skills and deeper learning dispositions, with demonstrable gains in learning and teaching (e.g., improvements in teacher-student relationships, as well as heightened critical reading engagement levels and skills.

1.2 My Groupwork Buddy

My Groupwork Buddy (MGB) aims to develop and integrate a techno-pedagogical system to encourage the 21CC of collaboration amongst students and support teacher pedagogical practices of facilitating student team projects. The project leverages existing collaborative inquiry tasks in the Secondary school curriculum and co-designs with teachers and policy officers to help students learn more about their personal teamwork competency and become more purposeful in their teamwork and learning.

MGB is guided by the “Team and Self Diagnostic Learning” (TSDL) framework of which key informing pedagogies are experiential learning, collaborative learning, and the learning analytics process model (Koh, Hong, & Tan, 2018; Verbert et al., 2013). TSDL is a process of developing students’ teamwork competency and comprises four stages: team-based concrete experience, self and team awareness building, self and team reflection and sensemaking, and self and team growth and change (See Figure 3).

Figure 3. The team and self diagnostic learning pedagogical framework of MGB

In brief, students’ begin with team-based concrete experiences namely, the collaborative inquiry task in any form, face-to-face or online. Following which students’ are made aware of their teamwork competency, primarily through self and peer ratings on a visual analytic in MGB. Next, students are guided to reflect on their team process and actions through the reflection space on MGB, with the teachers facilitating. Goal-setting and future-oriented questions are strategies in the process.

MGB functions includes an online synchronous chat, lesson content pages, self and peer ratings of teamwork competency, a teamwork visual analytic, a personal and team reflection space and status checks. The teamwork competency is based on a domain-neutral measure of teamwork comprising the dimensions of coordination, mutual performance monitoring, constructive conflict, and, team emotional support.

MGB was integrated into two partner schools’ normal curriculum subject (project work/design and technology) that employed collaborative inquiry tasks for 13-14 year olds. Trial 1 was completed in 2016 in 2 schools x 2 classes of 40 students each and 3 teachers. Trial 2 involved 2 schools x 4 classes of 40 students each and 8 teachers.

2 PEDAGOGICAL PROMISES AND PITFALLS

2.1 Potential of LA for enhancing learning outcomes and teaching practices

These two formative LA projects represent innovations that develop 21CC and productive learning dispositions to prepare future-ready students in Singapore. Research evidence generated from these design-based projects indicate the following benefits to learning and teaching:

- making visible 21CC learning processes and outcomes in a dynamic manner;
- cultivating connective literacy and other 21C literacies, such as collaboration, creativity, criticality among students;
- fostering greater self-awareness, reflective and self-regulatory learning dispositions;
- empowering autonomy and relevance as a critical feature of learning;
- enhancing learning motivation and engagement;
- enabling teachers' real-time formative monitoring of students' participation and progress, so as to provide more meaningful formative feedback;
- facilitating teachers' learning from the process of practitioner inquiry and adaptive expertise around data-informed pedagogy.

2.2 Potential pitfalls: LA visualisations a double-edged sword

The experienced pedagogical promises of LA notwithstanding, a number of pedagogical complexities and dilemmas were drawn from the user accounts of a critical stakeholder group—the learners. Specifically, the polarizing nature of peer-referenced LA dashboard visualizations was foregrounded. On one hand, many students found peer-referenced LA visualisations (i.e., visualisations comparing oneself to other peers) desirable for stimulating healthy competition and game-like learning. At the same time, there were strong opposing views that pointed to the adverse affective impact of such socially transparent peer comparisons, warning that such LA visualisations served to demoralize, pressurize, and trigger complacency in learners.

We provide a more comprehensive discussion of these pedagogical promises and pitfalls in our paper "Learner dashboards a double-edged sword? Students' sense-making of a collaborative critical reading and learning analytics environment for fostering 21st century literacies" available here (Tan, Koh, Jonathan and Yang, 2016).

3 CONCLUDING REMARKS: PEDAGOGICAL COMPLEXITIES AND ECOLOGICAL PERSPECTIVES OF LA IN SCHOOLS

Pedagogical complexities such as those outlined above foregrounds the need for systematic research on the impact of LA visualisations and dashboard designs on teaching and learning from user perspectives. Privileging the student and teacher voices, as critical stakeholder groups is paramount. Careful iterations of LA user-interface designs underpinned by validated pedagogical theories and learning principles is imperative. To this end, in a recent paper “Learning analytics in diverse educational contexts: Situating possibilities, paradoxes and pathways through an ecological pedagogic lens” available here (Tan and Koh, 2017), we argue for the need to situate LA as a pedagogical practice, and to understand pedagogy as an ecological sociocultural phenomenon.

Learning, and by implication LA, cannot be divorced from pedagogy. Rather, learning and pedagogy are conjoined endeavours that need to be theorised, understood, studied and designed in concert. Pedagogy—understood as the specific and cumulative relationships and interactions among learners, educators, the content and the environment—necessitates an ecological approach to designing LA as an important pathway forward for understanding and optimizing learning and teaching in schools.

ACKNOWLEDGEMENTS

This paper refers to data and analysis from the projects NRF2013-EDU001-EL019, NRF2015-EDU001-IHL08 and NRF2015-EDU001-IHL09 funded by the Singapore National Research Foundation, eduLab Research Program. The views expressed in this paper are the authors’ and do not necessarily represent the views of the National Institute of Education.

REFERENCES

Koh, E., Hong, H., & Tan, J. P. L. (2018). Formatively assessing teamwork in technology-enabled twenty-first century classrooms: exploratory findings of a teamwork awareness programme in Singapore. *Asia Pacific Jnl Education.* https://doi.org/10.1080/02188791.2018.1423952

Paul, R. & Elder, L. (2001). Critical Thinking: Tools for taking charge of your learning and your life. Upper Saddle River, NJ: Prentice Hall.

Tan, J. P-L., & Koh, E. (2017). Situating learning analytics pedagogically: Towards an ecological lens. Learning: Research and Practice, 3(1), 1-11. https://doi.org/10.1080/23735082.2017.1305661

Tan, J. P-L., Choo, S. S., Kang, T., & Liem, G. A. D. (2017). Educating for twenty-first century competencies and future-ready learners: research perspectives from Singapore. Asia Pacific Journal of Education, 37(4), 425-436. doi:10.1080/02188791.2017.1405475

Tan, J. P-L., Caleon, I.S., Jonathan, C.R., & Koh, E. (2014). A dialogic framework for assessing collective creativity in computer-supported collaborative problem-solving tasks. Research and Practice in Technology Enhanced Learning, 9(3), 411-437.

Tan, J. P-L., Koh, E., Jonathan, C. R., & Yang, S. (2017). Learner Dashboards a Double-Edged Sword? Students’ Sense-Making of a Collaborative Critical Reading and Learning Analytics Environment for Fostering 21st Century Literacies. Journal of Learning Analytics, 4(1), 117-140. http://dx.doi.org/10.18608/jla.2017.41.7

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. American Behavioral Scientist, 57(10), 1500-1509.

Mapping the Data Landscape in a Secondary School

David Ferguson
University of Technology Sydney
david.ferguson@student.uts.edu.au

ABSTRACT: The collection and analysis of data in secondary schools to support teaching and learning, leadership and decision making is becoming increasingly important. Ethical considerations require us to use such data to benefit learners and teachers, which compels users to understand the contexts in which data is collected and used. In partnership with a secondary school in Sydney Australia, a project was undertaken to map the school's data landscape. A diverse set of contexts including ethics, legislation, teaching and learning, decision-making and planning were identified in addition to the data sources and stakeholders producing and consuming data. A high-level data landscape was mapped and ideas for how to extend the work were shared with the school.

Keywords: learning analytics, secondary schools, data ethics, data context

1 BACKGROUND – WHY UNDERSTAND THE DATA LANDSCAPE?

As the collection and analysis of diverse types of data in schools increases and data is being used to support teaching and learning, leadership and decision making, it is important to understand the 'data landscape' within a school. A student project was undertaken with a NSW secondary school in partnership with the University of Technology Sydney that reviewed and documented the data landscape, along with the contexts in which data is or might be considered.

The data landscape can be thought of as a representation or map of the data currently available with the (school) environment. Prominent data sources, flows and uses of data within the school should stand out via this mapping, which aimed to use a wholistic view to draw attention to opportunities where it might be possible to improve or extend the use of data.

For the school involved in the project, the importance of data for teaching and learning is reflected in the school plan which included outcomes such as increasing the staff's ability to "use learning analytics to inform their teaching practice" as part of a collaborative and connected community of future focused learners, as well as to help "the school leadership team build the collective capacity of the staff and school community to use data to inform strategic school improvement effort".

2 METHODOLOGY

The project methodology included a range of information gathering approaches including conducting interviews with school contacts to review current report usage; identifying and interviewing other teachers and experts working with learning analytics in the NSW Secondary School system; seeking input from academics working in learning analytics at secondary & tertiary levels; and, reviewing and synthesising literature and departmental documentation. In all cases effort was made to translate information into a context relevant to learning analytics and data science at the school.

3 WHAT ARE THE DATA LANDSCAPE CONTEXTS?

Data contexts within secondary schools can take many forms. Some contexts are imposed externally through regulation while other contexts document the actors involved in producing and using data. In addition, there are ethical and strategic contexts which can support beneficial data use and decision-making.

3.1 Legislation and Standards

A range of legislation and standards exists for managing data in the public education context. The key pieces of legislation are the Education Act (NSW, 1990), Privacy and Personal Information Protection Act (NSW, 1998) and Health Records and Information Privacy Act (NSW, 2002). The NSW Department of Education has produced a Privacy Management Plan (2014) that notes exclusions in relation to the collection, retention and access to information to be considered alongside the application of the Information Protection Principles (2014a) and Health Privacy Principles (2014b).

3.2 Policies, Frameworks & Guidelines

The Centre for Education Statistics and Evaluation (CESE, 2017), is a NSW Department of Education body created in 2012 to improve the effectiveness, efficiency and accountability of education in New South Wales. CESE has published several data related guidelines for schools to use, including the collection and analysis of data, and the use of surveys which are available for schools to consult their students and parents.

3.3 Stakeholders

Five stakeholder groups were identified: Students; Teaching Staff; Executive Staff; Parents; and the Centre for Education Statistics and Evaluation (CESE). Each of these groups has one or more roles in producing, analysing and using data within the school landscape, or as part of the broader school system. Each group has its own roles and responsibilities, for example as a parent of the school, privacy legislation precluded me from viewing individual student level data, though this wasn't an issue for this project as we were considering the higher-level landscape.

3.4 Ethical use of data

Data ethics has been described as "a new branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use) … in order to formulate and support morally good solutions" (Floridi & Taddeo, 2016, p. 1). In the learning analytics context, it is important that any data science is done in a way that supports the students, their learning and the teachers who support and encourage that learning, without necessarily limiting the opportunity for providing value by being too restrictive.

Ferguson et al (2016) describe a number of types of ethical challenge in the context of learning analytics. The first challenge is described as there being a 'duty to act' if learner success is worth seeking (2016, p. 8). This challenge requires us to use data to benefit learners and teachers, rather than seeing ethical considerations as a limitation or something that stops is collecting and using data. This approach aligns with the stated school aim to "empower individual learners". Meeting this challenge will require efforts to ensure information is collected accurately and in a timely manner, that analysis is valid and comprehensible to those who use it and is presented in a way that supports teaching and learning.

3.5 Teaching and Learning

The school has been considering their teaching and learning in the context of the *4Cs* as described by Jefferson & Anderson in their (2017) book: *Creativity, Critical Reflection, Communication and Collaboration*. In this context, formative assessment of the 4Cs is likely to take a different form to conventional high stakes summative assessments, in order to work within a framework which is interdisciplinary, multidisciplinary, and transdisciplinary. The aim is "not to replace domain-specific knowledge – rather integrate the knowledge with broader skills students require for their 21st century lives" (Jefferson & Anderson, 2017, p. 175). Teachers are exhorted to "create connected tasks that assess Creativity, Critical Reflection, Communication, Collaboration as well as knowledge, wisdom & new understanding" (2017, p. 175). To measure achievement in ways such as this, there will be a need to create measurement systems and processes which facilitate the capture and collation of cross-department observations, and which ensure that each criterion is aligned to the 4Cs learning outcomes in addition to knowledge outcomes. As a parent, I recognise that consideration will also need to be given to how this new, broader assessment teaching and assessment framework (and the reasons for them) is communicated to, and understood by, parents in order to ensure broad support for the approach and inclusion of new metrics in reporting. The school's aspirations to develop learning analytics of this sort are aligned with other efforts nationally and internationally to develop new kinds of analytics for 21st century competencies (Buckingham Shum and Crick, 2016).

3.6 Planning and Decision Making

One of the stated aims of CESE is to promote the use of data to support decision making in education delivery, reflecting a theme in the School Excellence Framework and seen as an outcome in the school plan. In this context the capability to design programs which can be effectively evaluated is important, leading to evidence-based planning and decision-making.

In addition to teaching and learning based data sources, the availability and use of Learning Management and Business Reform (LMBR) data should be reviewed in this context. This data can be used to 'tell a story' to the broader school community as well as playing its role in decision making and school administration.

4 LANDSCAPE

4.1 Identified landscape components

Figure 1 presents a high-level representation of the key components of the Data Landscape. One of the limitations of this project was the inability to drill down to more detailed data (fields, flows and linkages) due to the conflict between student privacy and the author being a parent of the school.

With the appropriate privacy controls, a more detailed view of specific data types and flows could be uncovered and analysed, particularly between external assessment data, innovative programs, assessment outcome reporting and the third party 'Sentral' system.

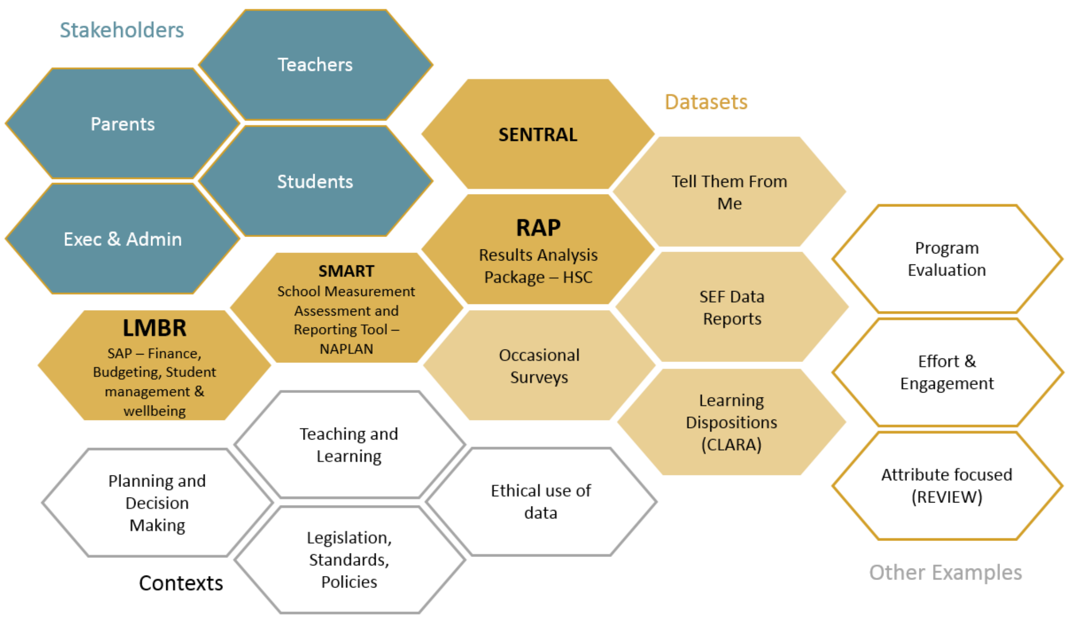


Figure 1: Data Landscape

4.2 Examples of innovative programs

A review of literature as well as interviews with secondary and tertiary practitioners identified examples of innovative approaches to learning data and analytics. Examples include the REVIEW system which is being piloted in a research project to test the viability of a focus on assessing capability development in high schools (Thompson 2017) and another where Nagy describes an approach and tool for tracking student effort (2016)

4.3 Identified gaps and opportunities

Despite the limitations of this project, a range of gaps and opportunities were identified within the school and in the project approach. As a result of staff changes, data ownership and guardianship was unclear in some situations (for example with the Sentral system). Improvements could include identifying owners as well as champions for using data to support teaching & evaluation throughout the school. Such a change would also support the bringing together of data from the various examples of innovative programs which were discovered across the school and having resources capable of assisting with designing data collection as part of program evaluation.

Practice improvements in project planning and execution will improve future data landscape projects. A standard agreement covering privacy and confidentiality with both school and centralised data should be created for reuse in future work. The discovery and analysis process for this project took in the order of 12 weeks and can be improved by involving department leads earlier and scheduling interviews in blocks, allowing the project to be completed in a 6 to 8 week timeframe. A secondary benefit of restructuring the project would be to draw out the many and distributed examples of individual teach data collection and connect them together in order to analyse and develop their true potential for supporting teaching and learning.

5 CONCLUSION

Understanding the contexts within which a secondary school's data landscape can be described is important to provide a complete picture of the opportunities for learning analytics. With knowledge about existing data and opportunities, steps can be taken to fill gaps in program design and data capture and to align activities with school plans and goals. This brief report illustrates an emerging methodology for mapping data landscapes in schools, which I look forward to discussing and refining.

Acknowledgements: I would like to thank the Principal and executive team at my partner school for participating in interviews and providing feedback as part of the project. I would also like to thank Stacey Quince and Duncan Rintoul from NSW Department of Education who offered their time to share information on innovative teaching and learning analytics.

6 REFERENCES

Buckingham Shum, S. and Crick, R. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics*, 3, (2), 6-21. http://dx.doi.org/10.18608/jla.2016.32.2

Centre for Education Statistics and Evaluation. (2017). Centre for Education Statistics and Evaluation. Retrieved 25 August 2017, from https://www.cese.nsw.gov.au

Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest Editorial: Ethics and Privacy in Learning Analytics. *Journal of Learning Analytics*, *3*(1), 5–15. https://doi.org/10.18608/jla.2016.31.2

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Phil. Trans. R. Soc. A*, *374*(2083), 20160360. https://doi.org/10.1098/rsta.2016.0360

Jefferson, M., & Anderson, M. (2017). *Transforming schools: creativity, critical reflection, communication, collaboration*. London, UK: Bloomsbury Academic.

Nagy, R. P. (2016). Tracking and Visualising Student Effort: Evolution of a Practical Analytics Tool for Staff and Student Engagement. Journal of Learning Analytics, 3(2), 164–192. https://doi.org/10.18608/jla.2016.32.8

NSW. Education Act (1990). Retrieved from http://www.austlii.edu.au/cgi-bin/viewdb/au/legis/nsw/consol_act/ea1990104/

NSW. Privacy and Personal Information Protection Act (1998). Retrieved from http://www.austlii.edu.au/cgi-bin/viewdb/au/legis/nsw/consol_act/papipa1998464/

NSW. Health Records and Information Privacy Act (2002). Retrieved from http://www.austlii.edu.au/ cgi-bin/viewdoc/au/legis/nsw/consol_act/hraipa2002370/index.html

NSW Government Education & Communities. (2014, March). Privacy Management Plan for Department of Education and Communities. Retrieved from https://www.det.nsw.edu.au/media/downloads/reports_stats/privacy/privacy-management-plan-march-2014.pdf

The Information and Privacy Commission NSW. (2014a, July 16). Information protection principles (IPPs) - for agencies [Text]. Retrieved 15 September 2017, from http://www.ipc.nsw.gov.au/information-protection-principles-ipps-agencies

The Information and Privacy Commission NSW. (2014b, October 7). The Health Privacy Principles (HPPs) explained for members of the public [Text]. Retrieved 7 September 2017, from http://www.ipc.nsw.gov.au/health-privacy-principles-hpps-explained-members-public

Thompson, D. G. (2016). Marks Should Not Be the Focus of Assessment – But How Can Change Be Achieved? Journal of Learning Analytics, 3(2), 193–212. https://doi.org/10.18608/jla.2016.32.9

Workshops

The Fourth LAK Hackathon: Benefiting Learning through novel data sources, standards and infrastructure

Alan Mark Berg
ICTS,University of Amsterdam
a.m.berg@uva.nl

Adam Cooper
Tribal Group
Adam.Cooper@tribalgroup.com

Kirsty Kitto
University of Technology, Sydney
kirsty.kitto@uts.edu.au

Mathiew d'Aquin
National University of Ireland
mathieu.daquin.w@gmail.com

Gábor Kismihók
University of Amsterdam
G.Kismihok@uva.nl

Niall Sclater
Sclater Digital Ltd
Niall.Sclater@jisc.ac.ukl

Tanya Dorey-Elias
Thompson Rivers University
telias@tru.ca

Daniele Di Mitri
Open University of the Netherlands
daniele.dimitri@ou.nl

Scott Harrison
University of Siegen
scott.harrison@uni-siegen.de

Anne-Marie Scott
University of Edinburgh
Anne-Marie.Scott@ed.ac.uk

ABSTRACT: Welcome to the Fourth LAK Hackathon. In this event, we emphasise expanding possibilities for improving the learner's experience through combining of novel data sources and infrastructures. If you have a research question, data source, idea or a problem bring it to the hackathon. We encourage joining the event no matter your background. We wish to mix it up with practitioners and researchers working in multidisciplinary teams towards common objectives.

Keywords: Hackathon, Learner Centric, Collaborative Development, Infrastructure, xAPI

1 A NOTE FROM THE ORGANISERS

Each year, for the last four years researchers and practitioners have run hackathons at the LAK conference. We have discussed many ideas, played about with the data, infrastructure, and learning practices. We have formed opinions, suggested strategies which radiate back to the LA research community as a whole. This year through the means of a Call For Proposal (CFP) we explicitly state research questions that we consider currently relevant and actionable, an agile research approach whose message we can amplify through evidence collection and message passing between events.

2 INTRODUCTION

In this fourth in the series of Hackathons held at LAK, we encourage a hands on approach to combining novel data sources in a realistic infrastructure for the benefit of Learning. The approach is multi-disciplinary, reviewed from all angles, self-organising, team building. Anyone is welcome to participate as long as they are motivated to be politely critical and work towards and expand common objectives. Please feel free to bring along your research questions, datasets and methodologies to the workshop for incorporation in the multidisciplinary activities.

3 BACKGROUND

Each year is based on the experiences gained from the previous years: The first hackathon, in **2015**, focused on the Apereo[1] Open Dashboard, with data sourced from an xAPI [3] Learning Record Store. It illustrated how the concept of an Open Learning Analytics architecture was developing but also shone light on some structural weaknesses: shortage of usable data for demonstration/development/quality-assurance/etc, and something of a gulf between the conceptions held by different stakeholders as to what a learning analytics dashboard would contain. Subsequent work by workshop organizers has begun to develop repeatable methods for generating synthetic data to help address the first weakness [1]. The second has been the topic of ongoing research (*vide infra*).

The second hackathon, in **2016**, continued to explore the **practicalities** of Open Learning Analytics. Using Jisc's emerging Learning Analytics architecture [8] as a reference point, with some data generated using the synthetic data methods which the first hackathon stimulated, the participants in the hackathon: scrutinized Jisc's interoperability recipes, tested the interoperability of learning record stores, learning analytics processors, and dashboards, and assessed the learning analytics standards landscape. The hackathon had a lasting effect, with numerous improvements to Jisc's interoperability recipes, and a strong message from the LAK community in favor of the greater integration of emerging learning analytics standards – **xAPI and Caliper** – contributing to the cooperation of ADL and IMS from mid-2016.

1 https://www.apereo.org/content/about

In the third hackathon, we built upon three assets: previous workshops, research, and recently-developed software. The first comprised the previous two LAK hackathons, the 2015 LAK Workshop "Visual Aspects of Learning Analytics" [2], and the 2016 LAK Workshop "Data Literacy for Learning Analytics" [9]. We had set the scene for the workshop using recent research on actionable analytics [6], student feedback [4], and embedding learning analytics in pedagogic practice [5]. Finally, introducing Jisc's student app, which was piloted with students across the UK after extensive consultation and design activities.

4 Organisational details

We have historically scheduled this inclusive, ***open workshop*** over two days. For the first half day, there is a period of orientation and an introduction to the core themes and mundane details such as how to interact with the data and the infrastructure. Team building from the very start, we place the participants in the centre of activity, evolving the schedule based on their feedback and expressed objectives. We then divide into teams of 6-8 to fulfil specific missions. At the end of each day we discuss progress, lessons learned and next steps. At the end of the second day, we summarise and plan future follow on actions.

This year, we will ask on a voluntary basis for short submissions detailing research questions and associated datasets. We will use these as opportunities to seed the set of problems teams can work and collect as part of a final report. Furthermore, this LAK hackathon will innovate upon previous organisational structures by hosting pre-hackathon events at the University of Technology Sydney's (UTS) Connected Intelligence Centre (CIC). UTS runs data and analytics programs. The pre-events will expose the audience to the student facing analytics, helping them to think about how they would like the area to evolve. Building their student facing LA solutions as a response to many different challenges set by the hackathon organisers (e.g. how to improve graduate employability, pathways to expertise, social connectedness). We will investigate sponsorship so that the winners of these pre-events can attend the main LAK hackathon and continue with their work on the global stage.

Logistics: We will need a large room (50 participants) with good internet connection, Beamer, tables for group exercises and if possible stationery such as sticky notes and pens. We assume that beverages are available and lunch for both days. A nice to have is a couple of small run off rooms for team's to find a quiet space. The organisers will provide an online presence, realistic LA infrastructure, seed datasets, GIT with slack and twitter channels.

5 Objectives and Outcomes

The principal aim of the hackathon is to enable multi-disciplinary thinking over key open challenges in Learning Analytics based on a problem-oriented, pragmatic approach. In line with the traditional definition of a hackathon, the expected outcomes of the event are the identification and initial (concrete, technical) pilot implementations of prototype tools/systems/data/studies which arise from the synthesis of educational technology, software development, and data science perspectives. As for previous events, the hackathon will generate a repository of code, sample data, screenshots, slides etc., from the activity of participants. An important intangible outcome will be an improved understanding of the different kinds of expert present about what is both desirable and technically-feasible.

While we welcome research questions, challenges, tools and data from all participants, we expect to emphasise the following topics which the organisers feel focus particularly on user-centred learning analytics:

Personal analytics supporting self-directed learning: Much of the data supporting the analysis of learning experiences are generated by the learners, and can be used to their benefit. Here, we are looking to understand what “learning analytics for the learner” can mean, especially in the context of self-directed learning using web platforms. The AFEL project has developed tools to collect such data from the learner’s browser and social media accounts, creating data spaces of online activities around the learner. Through the Hackathon, we hope to enable a greater understanding and initial practical solutions on the way those data spaces could be used to support learners in improving their own experience of using the Web for learning.

Goal setting and analytics: Goal setting is a potent tool to enhance the performance of individuals. However, we rarely use in education. With the help of learning analytics now, it is possible to set a goal and monitor those goals over time using data. This field is a novel and a vaguely explored area that gives a lot of room for creativity and development. We will build on the LAK16 Goal setting workshop [10] and available open source applications.

Playgrounds for data literacy: An emerging challenge for LA solutions concerns the lack of data literacy in both the academic and student populations. As we create more data and analytic models, can the people using it understand what it means? In alignment with the emerging field of critical data studies, there is an increasing need to develop an awareness among our students of the potential uses of their data and the possible consequences [11], including the development of tools that support this work. We first proposed the concept of a data playground in the 2017 Hackathon; we will return to this idea in the 2018 Hackathon to try to translate some of the initial ideas into demonstrable outputs.

Furthermore, we will continue to emphasise and develop the following enabling objectives:

Student facing Open APIs: While we are increasingly providing LA solutions for students, a significant opportunity arises to investigate the way in which universities and other open government services can enrich and expose their data stores and analytics, thus fostering the rapid development of student facing solutions. Institutions are increasingly moving towards an API based architecture which will add flexibility to their core IT infrastructure, a situation that offers many opportunities for rapid innovation and development of solutions by the people who will use them (rather than by external providers). At the same time, many universities host a pool of highly motivated students with fundamental ideas about how to improve the student experience by offering innovative new IT services that are beyond core business. This challenge is to facilitate such access pathways for systems built using the core infrastructure provided by official university data warehousing situations. We will investigate Security, data format, mapping and other core properties.

Infrastructure-integrated approaches for the joint exploitation of distributed data sources, including synthetic data: While previous events have focused on the exploitation of data from single sources, including the synthetic generation of data from those sources and their exploitation in visual analytics methods, a key aspect here regards the integration of the increasing number of data sources. From which we can assess the learning activities, learning environment and learners' experiences. Here, we will be looking at profiles and technical approaches to enable the joined-up use of data coming as much from institutional systems, as from the learners' other platforms (e.g. social media) [12]. A specific emphasis here will include the ability to generate synthetic, homogeneous and exploitable data from the kind of sources manipulated within the AFEL and Jisc projects.[2]

Analytics beyond user-computer interaction data. We will try to move from the datasets of user-to-computer interactions (the actor-verb-object paradigm) to focus more on the user-to-world interactions which we can elicit with motoric and physiological sensors (the sensor-sample-value paradigm). Multimodal datasets collected from practical and workplace learning experiences will be made available for analysis. These learning experiences are composed of atomic actions defined by an Experience API statement (e.g. the "expert pulls lever"), which will point to a list of multimodal sensor recordings. Some research questions associated with this type arise including: what are the essential features that we can extract from sensor data streams? Which data analysis and techniques are suitable? At the input level, how can we compare two or more action executions What is the most efficient way to integrate sensor data with xAPI? What is a meaningful visualisation?

[2] http://afel-project.eu

6 REFERENCES

[1] Berg, A.M. et al. 2016. The Role of a Reference Synthetic Data Generator within the Field of Learning Analytics. *Journal of Learning Analytics*. 3, 1 (2016), 107–128.

[2] Duval, E. et al. eds. 2015. VISLA 2015, Visual Aspects of Learning Analytics. *CEUR Workshop Proceedings* (2015).

[3] Experience API v1.0.1: 2013. *http://www.adlnet.gov/wp-content/uploads/2013/10/xAPI_v1.0.1-2013-10-01.pdf*.

[4] Khan, I. and Pardo, A. 2016. Data2U. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge – LAK '16* (New York, New York, USA, 2016), 249–253.

[5] Kitto, K. et al. 2016. Incorporating student-facing learning analytics into pedagogical practice. *Proceedings of the Annual ASCILITE Conference* (2016).

[6] Pardo, A. et al. 2016. Generating actionable predictive models of academic performance. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge – LAK '16* (New York, New York, USA, 2016), 474–478.

[7] Roll, I. and Winne, P.H. 2015. Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*. 2, 1 (2015), 7–12.

[8] Sclater, N. et al. 2015. Developing an open architecture for learning analytics. *EUNIS Journal of Higher Education*. (2015).

[9] Wolff, A. et al. 2016. Data literacy for learning analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge – LAK '16* (New York, New York, USA, 2016), 500–501.

[10] Mol, S., Kobayashi, V., Kismihók, G., & Zhao, C. (2016). Learning Through Goal Setting. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 512–513). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883859

[11] Pangrazio, Luci, and Neil Selwyn. (2017, September). "'My Data, My Bad ...': Young People's Personal Data Understandings and (Counter)Practices." *Proceedings of the 8th International Conference on Social Media & Society - #SMSociety17.* (pp. 1–5). ACM.

[12] Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016, April). Recipe for success: lessons learnt from using xAPI within the connected learning analytics toolkit. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 378-382). ACM.

Optimising E-portfolio's through the means of xAPI and Entity Extraction of Job Advertisements.

Alan Mark Berg
ICTS, University of Amsterdam
a.m.berg@uva.nl

Gábor Kismihók
Education and Labour Analytics Research group , TIB Hannover
Gabor.Kismihok@tib.eu

ABSTRACT: Within this hackathon theme, we explore the optimisation of e-portfolio's: Firstly, how entity extraction of authentic tasks from Vacancy data eases the pain of adoption and improves the learner's context. Secondly, how the standardising of xAPI e-portfolio profiles enable us to obtain scalable datasets for predictive models.

Keywords: E-Portfolio, Job Market Intelligence, Barriers to Adoption, Vacancy mining, Big Data, Entity Extraction

1 INTRODUCTION

Highly scalable e-portfolio systems exist including the open source systems such as Mahara (Gerbic & Maher, 2008), the next generation Open Portfolio System (Cambridge et al., 2008) and it's successor Karatu[1]. However, studies have consistently reported a negative perception of learners to e-portfolios (Rahayu & Sensuse, 2015). Issues include the difficulty of use for students, teachers and mentors, the quality of mentorship and the stability and complexity of the underlying system. In this hackathon, we explore how to make the student, teachers and mentors life easier through the application of Learning Analytics (LA) and Job Market Intelligence (JMI).

E-portfolio's have a potential to intersect with Learning Analytics, for example, Aguiar et al. (2014) showed how to use the activity of learners in e-portfolios to significantly improve the prediction of student dropout. Enabling e-Portfolio's through the xAPI protocol is a means of standardising the capture of the student digital trace and through this means eases the comparison between predictive models across organisations and encapsulation of learning moments, etc. Beckers, Dolmans, and Van Merriënboer (2016) suggested four central e-portfolio related themes: 1) Mandated as a dossier, which details

[1] http://karutaproject.org

achievements that an employer can hold against a standard. 2) The reflective e-portfolio, which is a voluntary version of the dossier and involves self-assessment on how the learner views themselves. This type of e-portfolio is often used when arguing for promotion. 3) The training e-portfolio, which aims at keeping track of learning, especially for employers and 4) the voluntary version is the personal development e-portfolio. We ask: *How do we capture these different usages through xAPI profiles and reuse the traces captured within predictive models?*

The authors acknowledge that E-portfolios are deployed for many different purposes and that scenario's 1) Detailing achievement and 3) keeping track of learning is probably the most straightforward situation's to define within xAPI statements/profiles. However, to aid in expanding the range we intend a) to provide a real E-portfolio system so the audience can experiment with the realistically structuring usage of portfolio's and b) offer a simple synthetic data generator (Berg, Mol, Kismihok, & Sclater, 2016) that is easily configurable to add new statements and then populate xAPI compliant applications. Through these means, we can quickly plug and play with the broader infrastructure provided for example by JISC. Another scenario is to plug and play with already existing recipes supplied for Goal setting (Berg, Scheffel, Drachsler, Ternier, & Specht, 2016). The audience can then experiment with streamlining interaction with Job Market Intelligence by encapsulating new events in the Job a market, such as newly discovered skills in a set of job adverts, again through the means of the xAPI standard.

In text mining, entity and feature extraction is a mature field with many decades of published research (Ramya et al., 2017; Nadeau & Sekine, 2007). Kobayashi et al. (2017a), provides a tutorial in the text analysis of job adverts, including code and data on how to perform extraction within the context of organisational research with specific reference to gathering the skills necessary for Nursing. Karakatsanis et al., 2017, conducted a broad scan of the job market for skill shortages based on the O-net database and Latent Semantic Indexing (LSI) and verified via crowdsourcing the accuracy of their approach and identified occupation clusters around skills, which are highly relevant within the collection of e-portfolios. Rahayu and colleagues (2017) reviewed recommendation systems that positively influence e-portfolio personalisation and concluded that researchers had shown the value of Hybrid and Collaborative Filtering, choosing the most relevant tasks together through online voting or tracking choice. One can easily imagine the content of Job adverts as a source of authentic tasks which feed the options for students as part of the e-portfolio recommendation process.

Based on the literature we suggest the following **research questions**:

> **RQ1:** *What is the definition(s) of an authentic task in the context of e-portfolio's?*
> **RQ2:** *How do we apply machine learning techniques to the extraction of authentic tasks?*
> **RQ3:** *How do we populate Karuta an open source e-portfolio system with authentic tasks?*
> **RQ4:** *What are the definitions of xAPI profiles for Job Market Intelligence enriched e-portfolio systems?*

RQ5: *Which variables captured by xAPI profiles describe the most variance in predictive models for student success?*

Logistics: Monsterboard has kindly donated 20,000 job advertisements. The authors have converted the data to Rdata and CSV format. We will also provide links and supporting help documentation for the installation of an opensource e-portfolio system with pointers for data scientists and programmers.

2 IMPACT

By populating a highly visable open source e-portfolio system with authentic tasks we signal to the market that entity extraction of job adverts is a viable approach to personalising the learner e-portfolio context. We expect this will lead to an accelerated market adoption, thus decreasing barriers to the acceptance by students and teachers of e-portfolio systems due to their at times unnecessary complexity and burden of use. Further, we extend the likelihood of further personalisations of e-portfolios.

Through the integration of e-portfolio systems with LA and JMI xAPI standards-based infrastructure, we increase the opportunity to compare and thus improve the accuracy and range of predictive models and the range of supporting services provided.

3 ACKNOWLEDGEMENTS

The author would like to acknowledge the critical feedback and support given by Assistant professor Stefan T Mol, and Professor Robin Boast and gratefully acknowledge the financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Program and Monsterboard.

4 REFERENCES

Aguiar, E., Ambrose, G. A., Chawla, N. V., Goodrich, V., & Brockman, J. (2014). Engagement vs Performance: Using Electronic Portfolios to Predict First Semester Engineering Student Persistence. *Journal of Learning Analytics*, *1*(3), 7–33. https://doi.org/10.18608/jla.2014.13.3

Berg, A., Scheffel, M., Drachsler, H., Ternier, S., & Specht, M. (2016). The dutch xAPI experience. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16* (pp. 544–545). New York, New York, USA: ACM Press. https://doi.org/10.1145/2883851.2883968

Berg, A., Mol, S., Kismihok, G., & Sclater, N. (2016). Scaling Learning Analytics: The practical application of synthetic data. In *Proceedings of the 9th European Distance and E-Learning Network Research Workshop* (p264-269).

Beckers, J., Dolmans, D., & Van Merriënboer, J. (2016). e-Portfolios enhancing students' self-directed learning: A systematic review of influencing factors. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.2528

Cambridge, D., Fernandez, L., Kahn, S., Kirkpatrick, J., & Smith, J. (2008). The impact of the open source portfolio on learning and assessment. *MERLOT Journal of Online Learning and Teaching*, *4*(4), 492-502.

Gerbic, P., & Maher, M. (2008). Collaborative self-study supporting new technology: The Mahara e-portfolio project. *Proceedings of ASCILITE Melbourne*, 320-324.

Karakatsanis, I., AlKhader, W., MacCrory, F., Alibasic, A., Omar, M. A., Aung, Z., & Woon, W. L. (2017). Data mining approach to monitoring the requirements of the job market: A case study. *Information Systems*, *65*, 1–6. https://doi.org/10.1016/j.is.2016.10.009

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2017a). Text Classification for Organizational Researchers. Organizational Research Methods, 1094428117719322. https://doi.org/10.1177/1094428117719322

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2017b). Text Mining in Organizational Research. Organizational Research Methods, 1094428117722619. http://dx.doi.org/10.1177/1094428117722619

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. Lingvisticae Investigationes, 30(1), 3-26. . https://doi.org/10.1075/li.30.1.03nad

Rahayu, P., Sensuse, D. I., Purwandari, B., Budi, I., Khalid, F., & Zulkarnaim, N. (2017). A Systematic Review of Recommender System for e-Portfolio Domain. In *Proceedings of the 5th International Conference on Information and Education Technology - ICIET '17* (pp. 21–26). New York, New York, USA: ACM Press. https://doi.org/10.1145/3029387.3029420

Ramya, R. S., Venugopal, K. R., Iyengar, S. S., & Patnaik, L. M. (2017). Feature Extraction and Duplicate Detection for Text Mining: A Survey. *Global Journal of Computer Science and Technology*, *16*(5). Retrieved from https://www.computerresearch.org/index.php/computer/article/view/1459

Hacking the Hackathon

Alan Mark Berg
ICTS, University of Amsterdam
a.m.berg@uva.nl

ABSTRACT: Hacking the hackathon: How do we increase the positive influence of the LAK hackathons and better embed into the broader context of the discussion between the Research and practitioner communities? How do we accelerate the trajectory of research impacting on the features and practices around Educational software? This submission explores the mission, organisation and help packages of hackathons to support persistence of effort and continuous exploration of research themes that have the immediate benefit to the enrichment of real systems.

Keywords: Hackathon, Community, Agile, Continuous

1 INTRODUCTION

The LAK hackathon[1](Cooper et al., 2016) is in its fourth incarnation. We have successfully supported; the development[2] of standard profiles for xAPI (Berg et al., 2016a), a discussion around synthetic data generation (Berg et al., 2016b), open source software such as the Apereo Learning Analytics Intitative[3] and the connectivist toolkit (Kitto et al, 2015) and delivered feedback on LA specific themes associated with the learners practices and dashboarding. JISC[4] and the Apereo Foundation[5] have in the past kindly provided infrastructure or logistic support and have in return had a timely delivery of feedback for example on the Jisc Student app[6] or their infrastructure (Sclater, Berg & Webb, 2015).

Each year we strive to improve the value of the hackathon to the research and practitioner communities and amplify the effect of the participants collective experiencess. For example, this year for the first time we have a call for proposal for research questions which, we will publish as part of the conference

[1] https://lakhackathon.wordpress.com
[2] https://lakhackathon.wordpress.com/about/history/
[3] https://www.apereo.org/communities/learning-analytics-initiative
[4] https://www.jisc.ac.uk
[5] https://www.apereo.org
[6] https://analytics.jiscinvolve.org/wp/2015/08/21/student-app-for-learning-analytics-functionality-and-wireframes

proceedings. We also have at least two pre-hackathon hackathons one in Holland and the other in Australia that will feed their collected experience into the LAK hackathon.

To continue with the improvement of these linked events we wish you to hack at the hackathon by answering the following interrelated research questions:

RQ1: Whom are our target audiences?
RQ2: What is the mission of a permanent space for LA hackathons?
RQ3: How do we fund the mission?
RQ4: How do we organise a permanent space?
RQ5: How do we scaleup any promising findings?
RQ6: How do we persist research questions and associated artifacts across hackathons?
RQ7: What is the definition of a hackathon support package?

Support and data: The hackathon website is permanent and is incrementally improved as new materials and idea’s become available. A GitHub location provides a safe coordinating space for the improvements of supporting material. GitHub is also a marshalling point for evidence collected across hackathons. Twitter's role is to advertise critical moments in the event, for example for submission dates for the Call For Proposals (CFP) and potentially we should incorporate communication via Twitter into the main hackathon event. Materials that are available from previous hackathons includes documentation, howto's, code and test plans and an example data generator. We currently expect a significant expansion of available datasets due to the nature of the CFP where we ask the authors to state their logistics and share explicitly.

Methodology: Currently, after a short introduction to potential themes and the supporting materials and practices, the audience organise themselves into groups. The organiser's of the hackathons divide themselves between the teams and support their activities. A couple of times a day the groups discuss together the progress they make, and a summary is prepared. Evidence and software artefacts are permanently stored on Github. This year we also have the opportunity to incorporate the materials generated from the pre-hackathons and defined by the submissions to the workshop. The presentations are expressing global themes that researchers in the field can collaborate and as such have the potential to focus broader efforts lasting longer than the hackathons themselves. We should consider reviewing the value of the digital traces and more determinedly collect data sources for re-use by other events.

2 IMPACT

When LA themed hackathons are successfully linked, we can develop a community adopted strategy guide which provides defined objectives, goals and milestones. When these goals are adequately funded, we stimulate the flow of actionable research towards adoption and accelerate the benefits to students, teachers and society.

3 ACKNOWLEDGEMENTS

The organisers gratefully acknowledge the logistical support of Jisc, the Apereo Foundation and SOLAR and the conference and especially the hackathon organisers some of whom we mention in this link[7] including Adam Cooper, Kirsty Kitto and Niall Sclater.

4 REFERENCES

Berg, A., Scheffel, M., Drachsler, H., Ternier, S., & Specht, M. (2016a). The dutch xAPI experience. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16* (pp. 544–545). New York, New York, USA: ACM Press. https://doi.org/10.1145/2883851.2883968

Berg, A. M., Mol, S. T., Kismihók, G., & Sclater, N. (2016b). The role of a reference synthetic data generator within the field of learning analytics. *Journal of Learning Analytics*, *3*, 107–128. https://doi.org/10.18608/jla.2016.31.7

Cooper, A., Berg, A., Sclater, N., Dorey-Elias, T., & Kitto, K. (2017). *LAK17 hackathon*. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17 (pp. 514–515). New York, New York, USA: ACM Press. https://doi.org/10.1145/3027385.3029435

Kitto, K., Cross, S., Waters, Z., & Lupton, M. (2015). Learning analytics beyond the LMS. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15* (pp. 11–15). New York, New York, USA: ACM Press. https://doi.org/10.1145/2723576.2723627

Sclater, N., Berg, A., & Webb, M. (2015). Developing an open architecture for learning analytics. *Proceedings of the EUNIS 2015 Congress*. https://doi.org/ISSN: 2409-1340

[7] https://lakhackathon.wordpress.com/organisers/

What role can Learning Analytics play in supporting university students to set goals for their own learning journey

Gábor Kismihók
TIB Hannover
Gabor.kismihok@tib.eu

Catherine Zhao
University of Sydney
Catherine.zhao@sydney.edu.au

ABSTRACT: This paper aims to create an opportunity for educational researchers and practitioners to explore how Learning Analytics (LA) can effectively support university students to set and achieve their learning goals. Goal-setting is thought to enable students to become skilled self-regulated learners, and in the long-term self-directed adult learners. This psychological and developmental process can be better understood and supported with the help of well-designed LA tools. LA has great potential in providing timely information on goal achievements. In particularly, open source frameworks and learning tools start to show some advantages in efficiency and greater degree of customisation. However, hurdles are present in the development of such tool sets. Disciplinary differences, institutional infrastructure, and organisational culture all play a catalyst role in the process. The authors provide two scenarios in the contexts of university library services and a university medical program. The first scenario raises the question of whether or not an API recipe a realistic objective, in light of the diversity of infrastructure settings and system adoptions across institutions, as well as differences in policy and governance. The second scenario attempts to investigate the meaning of goal-setting in an assessment driven educational experience, and how LA can help provide greater insight into students' psychological momentums along the learning journey. The authors hope the two scenarios to create an avenue for ongoing collaboration for new solutions to emerge.

Keywords: goal-setting, library services, medical education, API, open-source, Learning Analytics

1 BACKGROUND

As our societies are proceeding into the 21st century, technology and Big Data in education is not only radically changing the way we learn but also the strategies we choose to learn. As such, citizens' skills and competences to learn how to learn are identified as critical by researchers (Ramsden, 2003) and policy makers (EUR-LEX, 2017) alike. These skills and competences are crucial to success both in education and the workplace. From the learner's point, to become a self-regulated learner [Dale, Schunk & Zimmerman, 2013] has never been as important as it is today [Ames & Archer, 1988], when learning, regardless of whether it is formal or informal, needs to follow individual learning strategies. More focus should be directed towards individual student perspectives (Ferguson, 2012), and also to the social and technical context of learning (Shum & Ferguson, 2012).

Goal-setting, as an important dimension of Self-Regulated Learning (SRL) forehead stage, was the discussion focus in the Goal-setting workshop at Learning Analytics and Knowledge Conference (LAK) 2016. Participants suggested that GS should be an integral part of designing learning interventions (Wise et al., 2014). They also discussed the limited organizational uptake of GS, despite its demonstrated effects on study success. There is also evidence that learning analytics dashboards aid the visualization and internalization of learning goals and objectives (Scheffel et al., 2014; Verbert et al., 2014). Following the GS workshop and subsequent research work, this paper aims at continuing the conceptualization of GS and Learning Analytics (LA) interface [Mol et al., 2016]. The following scenarios in different university contexts serve this purpose by 1. Highlighting the potential of an 'API recipe' in library services that could help learners to formulate their learning and thinking strategies with real-time customized and holistic feedback; 2. Exploring possible avenues where learning analytics can cater for a given disciplinary education e.g. Medical education, where students are expected to achieve external goals set by the educators.

2 TWO SCENARIOS

Case I: Library

Discussions about library services in relation to LA are rare. Studies were looking at academic performance of students in relation to their activities in the libraries (Jantti at al. 2013; Soria et al. 2013). This limited interest in libraries is interesting, especially in the light that libraries are one of the places where actual learning takes place. Besides storing books and other printed and electronic content, libraries provide physical and virtual spaces for learners to interact with that content. Libraries also provide physical and virtual spaces to support interactions between learners (e.g. group discussion rooms) and also have vast amount of information about content searching behaviors.

For this reason, it might be very valuable to incorporate data coming from library systems into analytics of learning. However, as it was stated above, the available knowledge about pulling data from library system for LA is very limited. Therefore, the authors suggest that LAK18 hackathon participants try to 1, thinking about usable and feasible data sources about learning in libraries and 2, conceptualize possible requirements of integrating those datasets into LA. Ideally, these requirements should aim towards defining xAPI recipe(s) about activities learners are carrying out in libraries. Since a great number of educational and analytical software providers adopted successfully xAPI, delivering library services/learning related xAPI recipes might contribute to a faster uptake of library data in LA.

Case II: University medical programs (Medical education)

Medical education presents two unique characteristics in educational design. Firstly, unlike students enrolled in any other disciplinary programs, Medical students must go through the same 'channel' to become a doctor, that units of study are highly structured and mandatory. It represents a typical mastery learning experience, where students are expected to achieve goals set by educators. Secondly, assessment is traditionally ‘big’. In the case of the Sydney Medical Program, end-of-year assessment tasks are designed to test whether or not the student’s performance is satisfactory. The assessment outcome determines whether or not the student passes the whole year. In other words, assessment in medical education is the top driving force of student learning.

However, as the notion of assessment is evolving in the HE sector, programmatic assessment is gradually introduced into medical programs. In response to this change, assessment tasks are becoming more diverse. They can be designed to enable students to self-reflect, and self-evaluate their learning progress and direct their effort towards their personal learning goals. More importantly, students have some autonomy in making decisions for their learning, particularly in the increasingly popular self-directed online learning modules. In addition, the clinical schools that students are assigned to will benefit from in-time, detailed and accurate information to monitor the logistic aspects e.g. time-tabling, staffing, to provide personalized clinical learning experience for students, and to improve curriculum. As such, the value of LA in medical education increases. Information gathered by analytics tools has the potential to create an understanding of the holistic program experience for educators and students that allows them to be more proactive and to prioritize actions.

Based on the traditions of medical education and the directions it is moving toward, LA presents good opportunities to address two immediate problems – 1. How do educators provide timely and appropriate assessment feedback to students that they can carry forward into their learning plans; 2. What kinds of information is meaningful for educators and students with respect to the student’s personal development to become a skilled self-regulated learner.

3 RESEARCH QUESTIONS & CALL FOR COLLABORATION

Based on scenario 1 and 2, the authors aim to initiate a conversation in the Hackathon event at LAK18 by raising the following research questions

1: What library related (collected) datasets are useful when it comes to learning? How can those datasets can contribute to relevant LA methods? How library data can help learners to be better in SRL? And how library data can help teachers to design instructions/interventions?

2: What data is required to provide personalized assessment feedback to medical students? How should these sensitive data stored and managed in a university medical program? What’s the impact of policy and data governance on the development of open source tools for the aforementioned scenarios? An example of the Sydney Medical Program will be provided in the workshop as the baseline.

4 REFERENCES

Alyssa Wise, Yuting Zhao, and Simone Hausknecht. 2014. Learning analytics for online discussions: Embedded and extracted approaches. *Journal of Learning Analytics* 1, 2 (2014), 48–71.

Anne Conzemius and Jan O’Neill. 2009. The Power of SMART Goals: Using Goals to Improve Student Learning. Solution Tree Press.

Jantti, Margie and Brian Cox. Measuring the Value of Library Resources and Student Academic Performance through Relational Datasets. Evidence Based Library and Information Practice. Vol. 8, No. 2 (2013)

Soria, Krista M., Jan Fransen, and Shane Nackerud. Library Use and Undergraduate Student Outcomes: New Evidence for Students; Retention and Academic Success. portal: Libraries and the Academy. Vol 13, No. 2 (2013), pp. 147-164.

Carole Ames and Jennifer Archer. 1988. Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology* 80, 3 (1988), 260–267. DOI:https://doi.org/10.1037/0022-0663.80.3.260

Dale H. Schunk and Barry J. Zimmerman. 2012. Motivation and Self-Regulated Learning: Theory, Research, and Applications. Routledge.

EUR-Lex - c11090 - EN - EUR-Lex. Retrieved September 24, 2017 from http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:c11090

Katrien Verbert, Sten Govaerts, Erik Duval, Jose Luis Santos, Frans Van Assche, Gonzalo Parra, and Joris Klerkx. 2014. Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing* 18, 6 (2014), 1499–1514.

Maren Scheffel, Hendrik Drachsler, Slavi Stoyanov, and Marcus Specht. 2014. Quality indicators for learning analytics. *Journal of Educational Technology & Society* 17, 4 (2014), 117.

Mol, S.T., Kobayashi, V.B., Kismihók, G. and Zhao, C. 2016. Learning through goal setting. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 512–513

Paul R Pintrich. 2000. The role of goal orientation in self-regulated learning. *Handbook of self-regulation* 451, (2000), 451–502

Paul Ramsden. 2003. Learning to Teach in Higher Education. Routledge

Rebecca Ferguson. 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning* 4, 5–6 (2012), 304–317.

Simon Buckingham Shum and Rebecca Ferguson. 2012. Social learning analytics. *Journal of educational technology & society* 15, 3 (2012), 3.

Integrating privacy into an architecture for learning analytics

Author: George Ciordas-Hertel
German Institute for International Educational Research
ciordas@dipf.de

ABSTRACT: With the installation of the General Data Protection Regulation (GDPR) of the European Union each learning analytics architecture in place need to face specific privacy concerns. One of these concerns regards the right of the user to object to the collection of his personal data. In this hackathon a mechanism could be set in place which applies this right to a Learning Record Store (LRS).

Keywords: learning analytics, architecture, xAPI, privacy, GDPR

1. INTRODUCTION

The General Data Protection Regulation (GDPR) [2] of the European Union (EU) is intended to strengthen data protection for all European citizens. With this regulation the EU wants to give the people more control over how their personal data is used. The GDPR will be enforced in May 2018. Each learning analytics architecture in place will then need to face these privacy concerns. There are multiple challenges to face within such architecture [3]. A typical learning analytics system comprises multiple data sources which provide facts to a Learning Record Store (LRS). These facts could be xAPI statements resulting out of different types of activities the users do [1]. Considering "The right to object" of user to the collection of their personal data, the LRS needs to provide a mechanism to reject those facts.

2. CHALLENGE

A challenge the GDPR imposes on learning analytics architecture is that the facts which are getting pushed by multiple other systems need to be filtered with regard to the privacy settings of each individual user. As the GDPR describes a general right to object, giving the user a possibility to object more differential by activities, might allow us to still do analytics on parts of the personal data. Therefore a system with multiple components is necessary. One component is providing the privacy settings through a common API. Another component is the LRS which will be receiving and storing the facts. As you may see in figure 1, this LRS needs to be extended by a privacy guard which is responsible for the filtering based on the privacy settings. The purpose of this hackathon project is to design and implement a privacy filter mechanism for a learning analytics framework.

Research Questions:

- How to efficiently synchronize the privacy settings between software components?

- How to efficiently filter xAPI statements based on more or less generic privacy settings?

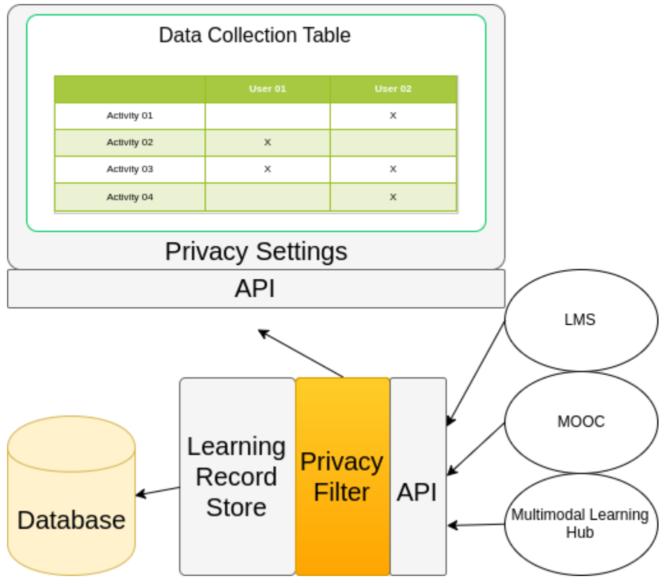


Figure 1: The privacy filter rejects facts arising from activities send by multiple sources based on the privacy settings of the users.

3. LOGISTICS

A system stub developed in Java EE 7 is available. The components are communicating with REST services sending JSON objects. The privacy component has a connection to a MySQL database. The LRS has a connection to a MongoDB. The whole system can be deployed on each development machine with a Docker setup. xAPI Statements could be send to the LRS from a separate system test component which afterward checks if the xAPI Statements where filtered out or not.

REFERENCES

[1] A. Berg and M. Scheffel and H. Drachsler and S. Ternier and M. Specht (2016). *Dutch Cooking with xAPI Recipes: The Good, the Bad, and the Consistent.* Proceedings of the 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT).

[2] European Commission. *REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL.* http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf

[3] Hoel, Tore and Griffiths, Dai and Chen, Weiqin (2017). *The Influence of Data Protection and Privacy Frameworks on the Design of Learning Analytics Systems.* Proceedings of the Seventh International Learning Analytics Knowledge Conference.

Multimodal challenge: analytics beyond user-computer interaction data

Daniele Di Mitri
Open University of The Netherlands
daniele.dimitri@ou.nl

Jan Schneider
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
schneider.jan@dipf.de

Marcus Specht
Open University of The Netherlands
marcus.specht@ou.nl

Hendrik Drachsler
Open University of The Netherlands – DIPF
hendrik.drachsler@ou.nl

ABSTRACT: this contribution describes one the challenges explored in the Fourth LAK Hackathon. This challenge aims at shifting the focus from learning situations which can be easily traced through user-computer interactions data and concentrate more on user-world interactions events, typical of co-located and practice-based learning experiences. This mission, pursued by the multimodal learning analytics (MMLA) community, seeks to bridge the gap between digital and physical learning spaces. The "multimodal" approach consists in combining learners' motoric actions with physiological responses and data about the learning contexts. These data can be collected through multiple wearable sensors and Internet of Things (IoT) devices. This Hackathon table will confront with three main challenges arising from the analysis and valorisation of multimodal datasets: 1) the data collection and storing, 2) the data annotation, 3) the data processing and exploitation. Some research questions which will be considered in this Hackathon challenge are the following: how to process the raw sensor data streams and extract relevant features? which data mining and machine learning techniques can be applied? how can we compare two action recordings? How to combine sensor data with Experience API (xAPI)? what are meaningful visualisations for these data?

Keywords: multimodal learning analytics, wearables, CrossMMLA, sensor-based learning

1 BACKGROUND

The Learning Analytics & Knowledge community has acknowledged the need for extending the analysis of learning to more diverse data sources going beyond the conventional online learning systems, MOOC events or student information systems. This need stems from the necessity of taking into account physical and co-located interactions which still represent the bulkiest set of learning activities.

Multimodal datasets can provide historical evidence and description of the learning process, i.e. the learner's behaviour and learning context. These data are collected automatically through wearable sensors, IoT devices and computer logs and therefore can capture only "what is visible" to some generic sensor. Such definition makes multimodal data conceptually separated by other human-driven qualitative interpretations like expert reports or teacher assessments. The latter are interpretations which describe dimensions unobservable with sensors, such as learning outcomes, cognitive aspects or affective states during learning. The analysis of multimodal data for learning has grown into a field of research called Multimodal Learning Analytics (Blikstein, 2013). These types of analytics seek to bridge complex behavioural patterns with learning theories (Worsley, 2014). In related work the multimodal approach was used in dialogic learning, during teacher-student discourses during lectures (D'mello et al., 2015); computer-supported collaborative learning during knowledge-sharing and group discussions (Martinez-maldonado et al., 2017; Schneider & Blikstein, 2015); and practice-based and open-ended learning tasks, when understanding and executing a practical learning tasks (Ochoa et al., 2013). The analysis of multimodal data in learning is a fairly new but steadily growing field of research which need support: The LAK community still lacks a programmatic approach for modelling the learning process and producing valuable feedback with multimodal data. In our understanding, this approach should clarify the collection, storage, analysis and exploitation of the multimodal data in a pragmatic and scalable manner, which can be adopted into real-life educational settings.

2 THE CHALLENGES

When describing and analysing learning with multimodal data, there still exist many open challenges (Blikstein & Worsley, 2016). For the LAK Hackathon, we identify three main challenges arising from the data-empowered feedback loop of multimodal data and learning analytics: 1) the data collection and storing; 2) the data annotation (or triangulation); 3) the data processing and exploitation.

2.1 Data collection and storing

The first step of the journey is the *data collection* and the data creation with new multimodal datasets. The sensors are most likely from different vendors and have different specifications and support, the approach used for data collection must be flexible and extensible to different sensors which collect data at different frequencies and formats. To address this challenge, we introduce the *LearningHub*, a software prototype whose purpose is to synchronise and fuse the different streams of multimodal data generated by the different sensor-applications while capturing a meaningful part of a learning task, that we call *Action Recording*. The *LearningHub* channels the data from multiple sensors and provides as output JSON files, which serialise and synchronise the values of the sensors for each sensor application. The JSON files allow to have multiple attributes with different time frequencies and formats; they provides also the logic to facilitate the action recording for storing and later retrieval. Research questions: 1.1) how to improve the *LearningHub*? 1.2) is the JSON serialisation of the Action Recoding the best approach for storing and retrieving? 1.3) how to link an Action Recording to an xAPI?

2.2 Data annotation

The *data annotation* challenge consists in finding a seamless and unobtrusive approach for labelling the learning process, i.e. triangulating the multimodal action recordings with evidence of the learning activities. To address this challenge, we propose another software prototype: the *Visual Inspection Tool (VIT)*. The VIT is useful for retrospectively analysing and annotating multimodal action recordings. The VIT allows to load multimodal datasets, plot them on a common time scale and triangulate them with video recordings of the learning activity. It allows to select a particular timeframe and annotate the multimodal data slice with a xAPI triplet, assigning an actor, a verb and an object. The VIT offers a human-computer interface which helps to deal with the complexity of multimodal data. Research questions: 2.1) how to best improve the VIT? 2.2) how to define and exchange xAPI vocabulary for multimodal activities?

2.3 Data processing and exploitation

The data processing step consists in extracting and aligning the relevant features from the "raw" multimodal data and making them suitable for exploitation, meaning by providing some personalisation benefits to the learner or the teacher. The data processing steps depend on the chosen exploitation strategy. For example, *light-weight feedback* can be generated through hardcoded rules; *historical reports* require different visualisations in that can be grouped into an analytics dashboard; *frequent patterns* or *predictions* require training either unsupervised or supervised models, store them into memory and use them to estimate the value or the class of a particular target attribute. Research questions: 3.1) how to process the raw sensor data streams and extract relevant features? 3.2) which data mining and machine learning techniques can be applied? 3.3) how can we compare two action recordings? 3.4) what are meaningful visualisations?

REFERENCES

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 102. http://doi.org/10.1145/2460296.2460316

Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining : using computational technologies to measure complex learning tasks, *7750*(x), 220–238. http://doi.org/http://dx.doi.org/10.18608/jla.2016.32.11

D'mello, S., Olney, A., Blanchard, N., Sun, X., Ward, B., Samei, B., & Kelly, S. (2015). Multimodal Capture of Teacher-Student Interactions for Automated Dialogic Analysis in Live Classrooms. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 557–566. http://doi.org/10.1145/2818346.2830602

Martinez-maldonado, R., Power, T., Hayes, C., Abdiprano, A., Vo, T., & Shum, S. B. (2017). Analytics Meet Patient Manikins : Challenges in an Authentic Small-Group Healthcare Simulation Classroom, 1–5. http://doi.org/10.1145/3027385.3027401

Ochoa, X., Chiluiza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI '13)*, 583–590.

Schneider, B., & Blikstein, P. (2015). Unraveling Students' Interaction Around a Tangible Interface using Multimodal Learning Analytics. *JEDM - Journal of Educational Data Mining*, *7*(3), 89–116.

Worsley, M. (2014). Multimodal learning analytics as a tool for bridging learning theory and complex learning behaviors. *3rd Multimodal Learning Analytics Workshop and Grand Challenges, MLA 2014*, 1–4.

A Learning Analytics Data Literacy Playground

Tanya Dorey-Elias
University of Calgary
eliasj@telus.net

Anne-Marie Scott
University of Edinburgh
anne-marie.scott@ed.ac.uk

ABSTRACT: This paper proposes a **Hackathon Research Question**. An emerging challenge for learning analytics solutions concerns the lack of data literacy in both the academic and student populations. As we create more data and analytical models, can the people using them understand what they mean? Do they understand the risks associated with sharing their data? In alignment with the emerging field of critical data studies, there is an increasing need to develop an awareness among our students of the potential uses of their data and the possible consequences including the development of tools that support this work. We first proposed the concept of a data playground in the 2017 Hackathon. For LAK18, we return to this idea and challenge participants to translate some of the initial ideas into demonstrable outputs.

Keywords: Data literacy, digital literacy, ethics, digital understanding

1 INTRODUCTION

An emerging challenge for learning analytics solutions concerns the lack of data literacy in both the academic and student populations. As we create more data and analytical models, can the people using them understand what they really mean and do they understand the risks associated with the ways their data is being used?

The LAK18 Conference asks:

- Which design processes involve learners, educators and other users effectively in the co-design of analytics tools?
- Which techniques are effective in assessing how end-users make sense of, interact with, and act on analytics feedback?
- In what ways can learning analytics systems be biased, and can they be made more transparent and accountable to different stakeholder groups?
- How are educational leaders creating the conditions for learning analytics systems to take root and grow?
- How strong is the evidence that the adoption of learning analytics benefits stakeholders?

(LAK18, 2017)

We argue that before these issues can be tackled within institutions and research projects, thought must be given to how to best build critical data literacy skills with collaborators and stakeholders. In alignment with the emerging field of critical data studies, there is an increasing need to develop an awareness amongst our students in particular of the potential uses of their data and the possible consequences (Pangrazio and Selwyn, 2017). More information about student perspectives of learning analytics needs to be gathered to better inform learning analytics implementations and algorithms need to be open to interrogation (Knox 2017a). Skills for working with and interrogating data itself need to be built, rather than focusing solely on outputs and visualisations (D'Ignazio and Barghava, 2016). Successful co-design of analytics tools, for example, requires that students have a baseline level of understanding as to what data about them are being collected and what they could be used for both within educational institutions and beyond. Underlying assumptions about trust, ethics and duty of care also need to be articulated, balanced and understood.

2 LEARNING ANALYTICS DATA LITERACY PLAYGROUND DESCRIPTION

We propose as a challenge for the LAK18 Hackathon the creation of a "Learning Analytics Data Literacy Playground." This follows on from, and builds upon thinking that emerged from the LAK'17 Hackathon (Dorey-Elias 2017). We envisage that this challenge would include the development of synthetic data sets that mirrors data gathered while a student including VLE data, student record data, attendance monitoring data, etc. Students may be able to optionally include social media feeds (their own, or potentially another synthetic data set). A user interface or set of tools that allows students to interrogate, experiment with, and build on this data would then be developed. The playground should have some basic structure or set of tasks in place that scaffold initial activities and provide an easy route into exploration. It should challenge assumptions and reveal potential insights from data that may not be immediately obvious. It should also collect insights from students through in-built feedback mechanisms such as asking questions at key points, or recording decisions made.

Students interrogation and exploration of the playground could in itself be a source of more insight. This should be considered in the design. For example, at the end of a session, the system could display a summary of activities completed for the student. This information might include how similar their interrogations were to those of other students, and allow them to generate some sort of a "receipt" for their activity, or alternately, allow them to delete all trace of their activity. The playground could also include insight into the legal basis for data collection and processing based on selection of a country profile.

The purpose of the playground should not be to scare, but rather to educate students. Interaction with the playground should be engaging, and playful if possible (see for example https://dataselfie.it/#/ or https://applymagicsauce.com/). It could be modelled on a "choose your own adventure" branching pathway; it could be an interactive dashboard on which many of the inputs and outputs can be calibrated by the end user. Exploring the interface design choice will be as important as the manipulation of data in terms of developing something that students can interact with, make sense of, and act upon.

This challenge has a broad scope and so the initiation activity will be for participants to identify use cases, and then agree a small number on which the rest of the Hackathon activity will be focused.

3 LEARNING ANALYTICS DATA LITERACY PLAYGROUND IMPACT

We see the Data Literacy Playground as having a range of potential uses including:

- To support learning activities in a wide range of academic courses related to digital and data literacies and digital cultures.
- As an orientation exercise at the start of any co-design activity within an organization.
- To improve student understanding as to the types of data typically gathered about them and how it might be used, including among services not associated with the academic institution.
- To build a better understanding of informed consent for the use of services.
- As a research tool to further gather insight into student attitudes to learning analytics.

We are excited about this opportunity to advance the development of a tool through which students would be empowered to “read, work with, analyze and argue with data as part of a larger inquiry process” (D’Ignazio and Barghava, 2016).

REFERENCES

D'Ignazio, C., & Bhargava, R. (2016). DataBasic: Design Principles, Tools and Activities for Data Literacy Learners. *The Journal of Community Informatics, 12*(3). Retrieved from http://ci-journal.net/index.php/ciej/article/view/1294/1229

Dorey-Elias, T. (2017). *Building a digital literacy playground – an emerging idea*. *Here to there*. Retrieved from https://heretothere.trubox.ca/building-a-digital-literacy-playground-an-emerging-idea/

Knox, J. (2017a). Data Power in Education: Exploring Critical Awareness with the “Learning Analytics Report Card”. *Television & New Media*, *18*(8), 734-752. doi: 10.1177/1527476417690029

Knox, J. (2017b*).* Playing with student data: The Learning Analytics Report Card (LARC). In S. Shehata, JP-L Tan (Eds.), *Practitioner Track Proceedings of the 7th International Learning Analytics & Knowledge Conference,* 43-49.

LAK18 (2017). General Call, The 8th International Conference on Learning Analytics & Knowledge Learning Analytics and Knowledge 2018 [web log post]. Retrieved from https://latte-analytics.sydney.edu.au/general-call/

Pangrazio, L., & Selwyn N. (2017). ‘My Data, My Bad ...’: Young People’s Personal Data Understandings and (Counter)Practices. *Proceedings of the 8th International Conference on Social Media & Society,* 1–5. doi: 10.1145/3097286.3097338

2nd CrossMMLA: Multimodal Learning Analytics Across Physical and Digital Spaces

Roberto Martinez-Maldonado
University of Technology, Sydney, Australia
roberto.martinez-maldonado@uts.edu.au

Vanessa Echeverria
University of Technology, Sydney, Australia Escuela
Superior Politécnica del Litoral, ESPOL, Ecuador
vanessa.i.echeverriabarzola@student.uts.edu.au

Luis P. Prieto
School of Educational Sciences, Tallin University, Estonia
lprisan@hotmail.com

Maria Jesus Rodriguez-Triana
École Polytechnique Fédérale de Lausanne & Tallinn University, Estonia
maria.rodriguez-triana@epfl.ch

Daniel Spikol
Malmö University, Sweden
daniel.spikol@mah.se

Mutlu Curukova and Manolis Mavrikis
UCL Knowledge Lab, United Kingdom
M.Cukurova@ucl.ac.uk, m.mavrikis@ucl.ac.uk

Xavier Ochoa
Escuela Superior Politécnica del Litoral, ESPOL, Ecuador
xavier@cti.espol.edu.ec

Marcelo Worsley
Northwestern University, United States
marcelo.worsley@northwestern.edu

ABSTRACT: Students' learning is ubiquitous. It happens wherever the learner is rather than being constrained to a specific physical or digital learning space (e.g. the classroom or the institutional LMS respectively). A critical question is: how to integrate and coordinate learning analytics to provide continued support to learning across physical and digital spaces? CrossMMLA is the successor to the Learning Analytics Across Spaces (CrossLAK) and

MultiModal Learning Analytics (MMLA) series of workshops that were merged in 2017 after successful cross-pollination between the two communities. Although it may be said that CrossLAK and MMLA perspectives follow different philosophical and practical approaches, they both share a common aim. This aim is: deploying learning analytics innovations that can be used across diverse authentic learning environments whilst learners feature various modalities of interaction or behaviour.

Keywords: Learning analytics, seamless learning, integration, multimodal

1 WORKSHOP BACKGROUND

1.1 Motivation

Educational research has revealed the pedagogical benefits of letting students experience different types of content, "real world" challenges, and physical and social interactions with educators or other learners (Delgado Kloos, Hernández-Leo, & Asensio-Pérez, 2012). This is partly because student's learning happens in situ, where the learner is (Sharples, M., & Roschelle, 2010). Learning is not necessarily constrained to a specific physical space (e.g. the classroom) or a digital environment (e.g. an institutional learning management system or a specific learning digital tool). Moreover, in practice, students commonly work outside the boundaries of the institutional learning system(s). This inherently blended nature of learning settings makes it essential to move beyond learning analytics that rely solely on a single data source (e.g. log files) or that focus only on the interactions that occur between learners and a specific system without considering the context of use. A critical question is: how to integrate and coordinate learning analytics to provide continued support to learning across physical and digital spaces?

CrossMMLA is the successor to the Learning Analytics Across Spaces (CrossLAK) and MultiModal Learning Analytics (MMLA) series of workshops that were merged in 2017 after successful cross-pollination and synergetic efforts between the two communities. CrossLAK and MMLA perspectives follow different philosophical and practical approaches. It may be said that CrossLAK follows a top-down approach, focusing on learning first and then on the analytics. First, it embraces the complexity of learning as an activity which is distributed across spaces, people, tools (both digital and physical) and time. Once the "learning problem" has been identified, a CrossLAK initiative would analyse the feasibility of using learning analytics to tackle such a problem. These analytics may be very simple (unimodal) or quite sophisticated (multimodal). Since the focus is on learning happening in authentic spaces, the philosophical intention is to apply analytics in-the-wild rather than in-the- lab.

By contrast, we can say that MMLA favours a bottom-up approach where the focus is on the analytics grounded by learning theory and practice. MMLA can provide insights into learning processes that happen across multiple contexts between people, devices and resources (both physical and digital), which often are hard to model and orchestrate (Scherer, Worsley & Morency, 2012; Worsley et al., 2015; Prieto et al., 2016; Ochoa et al. 2017). MMLA leverages the increasingly widespread availability of sensors and high-frequency data collection technologies to enrich the

existing data available. Using such technologies, in combination with machine learning and artificial intelligence techniques, a number of solutions can be offered to ubiquitous learning. Although, several MMLA projects have been conducted in-the-lab (see review in Ochoa, 2017), the intention of this joint workshop is for MMLA to also move into-the-wild.

Although CrossLAK and MMLA have some elements that distinguish them from each other, they both share the common aim of deploying learning analytics innovations that can be used across diverse authentic learning environments whilst learners feature various modalities of interaction or behaviour. LA researchers can now perform text, speech, handwriting, sketch, gesture, affective, neurophysical, or eye gaze analyses (Donnelly et al. 2016; Prieto et al., 2016). Collecting and understanding data from the everyday learning environments becomes increasingly challenging. However, pervasive and mobile technologies can be used to allow learners to get remote access to educational resources from different physical spaces (e.g. ubiquitous/mobile learning support) or to enrich their learning experiences in the classroom in ways that were not previously possible (e.g. face-to-face/blended learning support). This is creating new possibilities for learning analytics to provide continued support or a more holistic view of learning, moving beyond desktop-based learning resources.

Our aim as a joint CrossMMLA community is to make learning analytics relevant across, physical, digital, and blended learning environments while making the tools more accessible to the wider community. Therefore, researchers and practitioners need to address the larger frame of what is happening across the digital and physical space and between individuals, groups, and the entire class while balancing the data, collection, analysis and visualisation.

1.2 Contribution to LAK 2018

Aligned with LAK'18 interest on learning analytics adoption, in this workshop we want to pay special attention on how our CrossMMLA solutions may move out of the lab into the real world, reflecting not only on technical criteria, but also on the adoption and effectiveness in authentic settings. More concretely, the contributions (via attendance and paper submissions) should emphasise the considerations taken in the design, deployment and assessment of the proposals in order to demonstrate its feasibility and potential to be successfully and sustainably implemented.

1.3 Evidence of Interest and Previous Events

Current Special Issue call:

- UMUAI Journal: Special issue on Multimodal Learning Analytics and Personalised Support

Across Spaces [call for papers]

Some of the previous editions of MMLA and across spaces WS. Proceedings available.

- CrossMMLA17 @ ECTEL'17 - International Workshop on Multimodal Learning Analytics Across Spaces. This was one of the most-attended workshops in EC-TEL 2017, with more than 20 participants.
- MMLA17 @ LAK'17 - 6th Multimodal Learning Analytics Workshop [CEUR proceedings]
- CrossLAK17 @ LAK'17 - 2nd International Workshop on Learning Analytics Across Spaces
- UMAP '17 - Tutorial: Designing Cross-Space Learning Analytics and Personalised Support
- CrossLAK16 @ LAK'16 - International Workshop on Learning Analytics Across Spaces [CEUR proceedings]
- MMLA @ LAK'17 - Current and Future Multimodal Learning Analytics Challenge (6th edition)
- MMLA @ LAK'16 - Fifth Multimodal Learning Analytics Data Challenges
- AcrossSpaces @ EC-TEL'11 - Learning activities across physical and virtual spaces (AcrossSpaces) workshop

2 ORGANISATIONAL DETAILS

2.1 Type of Event and Schedule

The workshop is envisioned to be a 1 full day workshop divided in three parts:

1. The first part will be a **panel discussion** focused on intermediate constructs/indicators in CrossMMLA (a recurring topic that emerged in the last workshops).
2. The second part will be a **hands-on ideation task** focused on identifying critical systems, tools, standards in MMLA (e.g., towards a unified CrossMMLA stack).
3. The third part will be a **practical/hands-on MMLA training** using multimodal sensors (e.g. multimodal selfies, beacons, etc.) and tools to then explore/analyse the data.

Authors' papers will be presented as posters, with a poster madness session at the beginning of the day with some discussion around the posters during the breaks.

2.2 Participation

All workshop participants are encouraged to submit at least one short paper. There is no restriction in the number of papers submitted by the same author. The submission of a paper is not compulsory.

After the conference, full-length submissions resulting from the discussions will be published on CEUR (http://ceur-ws.org), and we are exploring the joint edition of future journal special issues on the topic.

2.3 Equipment required

We will require around 10 poster stands or a room where posters can be hanged on the walls using sticky tape. We would like to run the workshop in a room that allows us to configure tables in a way it allows small group activities (a collaboration room would be ideal). A projector and a whiteboard would be required.

3 OBJECTIVES AND OUTCOMES

One of the key aims of the workshop is to attract researchers (from diverse communities) to consider how multimodal learning analytics can be used across diverse learning environments. The intention is to gather interested parties in ubiquitous, mobile and/or face-to-face learning analytics with a focus on multimodal interaction.

REFERENCES

Delgado Kloos, C., Hernández-Leo, D., & Asensio-Pérez, J. I. (2012). Technology for learning across physical and virtual spaces: J. UCS special issue. *J Univers. Comput Sci. 2012; 18 (15): 2093-2096.*

Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., ... & D'Mello, S. K. (2016, July). Automatic teacher modeling from live classroom audio. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*(pp. 45-53). ACM.

Ochoa, X. (2017) Multimodal Learning Analytics. *Handbook of Learning Analytics,* Lang, C.; Siemens, G.; Wise, A. & Gašević, D. (Eds.). Society for Learning Analytics Research (SoLAR), 129-141

Ochoa, X., Worsley, M., Weibel, N., & Oviatt, S. (2016, April). Multimodal learning analytics data challenges. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 498-499). ACM.

Prieto, L. P., Sharma, K., Dillenbourg, P., & Rodríguez-Triana, M. J. (2016, April). Teaching analytics: towards automatic extraction of orchestration graphs using wearable sensors. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 148-157). ACM.

Scherer, S., Worsley, M., & Morency, L. P. (2012, October). 1st international workshop on multimodal learning analytics. In *Proceedings of the 14th ACM international conference on Multimodal interaction* (pp. 609-610). ACM.

Sharples, M., & Roschelle, J. (2010). Guest editorial: Special section on mobile and ubiquitous technologies for learning. *IEEE Transactions on Learning Technologies*, *3*(1), 4-6.

Worsley, M., Chiluiza, K., Grafsgaard, J. F., & Ochoa, X. (2015, November). 2015 Multimodal Learning and Analytics Grand Challenge. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 525-529). ACM

Learning false friends across contexts

Author(s): Victoria Abou-Khalil
Graduate School of Informatics, Kyoto University
v.aboukhalil@gmail.com

Author(s): Brendan Flanagan
Academic Center for Computing and Media studies, Kyoto University
flanagan.brendanjohn.4n@kyoto-u.ac.jp

Author(s): Hiroaki Ogata
Academic Center for Computing and Media studies, Kyoto University
hiroaki.ogata@gmail.com

ABSTRACT: False friends are words in two languages that look or sound similar but differ significantly in meaning in some or all contexts. False friends are confusing for language students and could result in frustration and communication problems. This paper proposes a method to diagnose and prevent false friends mistakes based on students' past learned words, current location and time. The proposed method uses records from the SCROLL system (System for Capturing and Reminding Of Learning Log) to analyze the previous activity of students. We assume that the past activity of a student can be used to predict the meaning intended by the student when looking up a polysemous word. The identification of the intended meaning in the student's current context is then used to provide the student with the appropriate translation, warnings and quizzes, possibly improving the learning process and avoiding false friends future mistakes.

Keywords: Learning Analytics, Ubiquitous learning, False Friends, Computer Supported Language Learning

1 INTRODUCTION

When learning a second language, students can take advantage of the vocabulary of their first language using cognates (Nation, 2003). Cognates are words that sound or look similar in the two languages, have similar meanings, and help students expand their vocabulary by playing the role of `true friends'.

However, we can all imagine the awkward situations that could arise if the French word promiscuité (lack of privacy, crowding) is interchanged with the English word *promiscuity* (i.e., having a lot of different sexual partners) in a sentence. We can also imagine flirting going wrong when a French speaker compliment their English crush using the word *formidable* (i.e., inspiring fear or respect through being impressively large or powerful), when what she actually meant is *formidable* (i.e., inspiring awe). We can understand how strange a Japanese speaker will sound when saying `My mansion is on the second floor', when what he had in mind was マンション (mansion, i.e., flat, apartment). The previous situations are faced by language learners and are caused by a tricky category of words: false friends.

False friends are words in two languages that look and sound similar, but differ significantly in meaning in some or all contexts. The degree of complexity of learning false friends depends on whether they are total false friends or partial false friends. Total false friends have completely different meanings in both languages (e.g.: Eng.: *Attend* (to be present); French: *Attendre* (wait)), whereas partial false friends are polysemous words, one of whose meanings is a false friend while others are true cognates (e.g.: Eng.: *Demand* (i.e., request made as of right); French: *Demander* (i.e., to ask; to be looking for; to demand). Figure 1 shows the types of `friends' that a student will encounter when learning the target language.

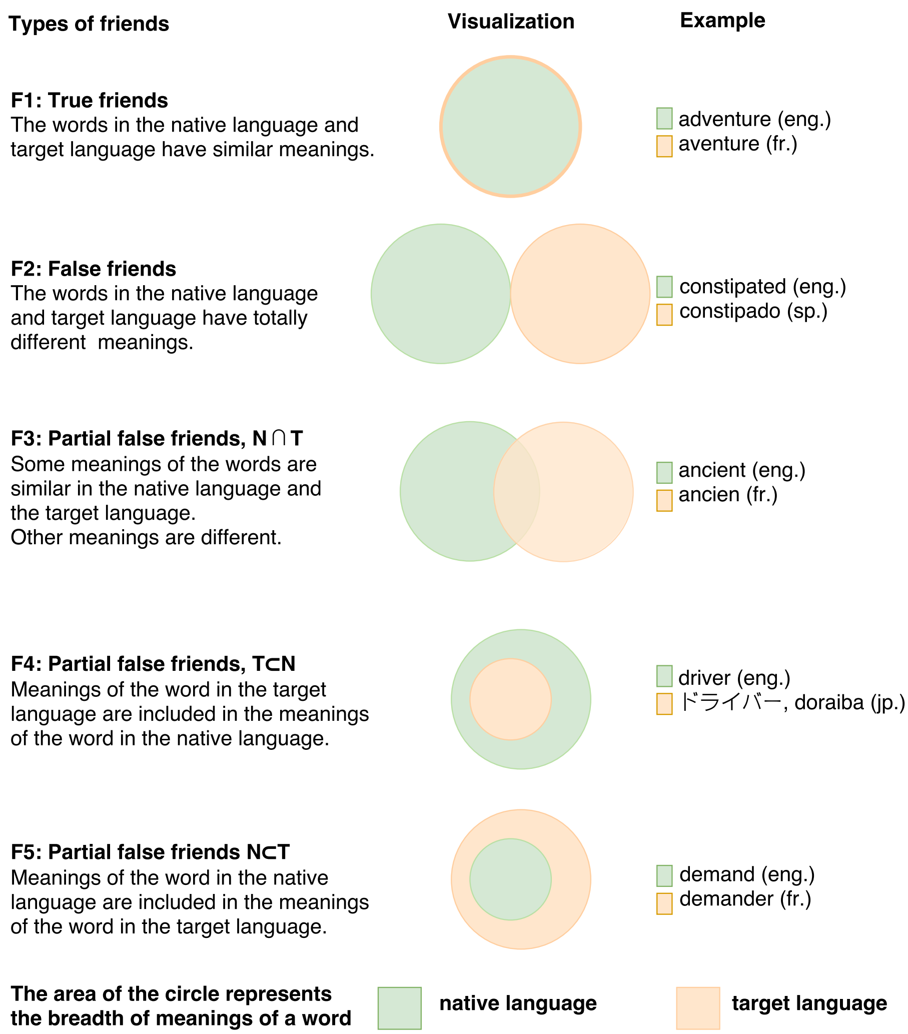


Figure 1: Types of false friends between a student's native language and the learned target language.

Depending on the context, partial false friends play the role of a true or a false friend. This paper proposes a method to diagnose and prevent false friends mistakes based on students' past learned vocabulary, current location and time. First, we analyze the factors that lead to false friends confusions while learning the target language. We also analyze the types of assimilation problems that could then arise, depending on the type of false friend. In order to prevent false friends confusions, we use the student's past learning logs (previous looked up words, locations, time) to predict the meaning intended by the student when looking up a polysemous words. The identification of the intended meaning is then used to provide the student with the appropriate translation and warnings, possibly improving the learning process and avoiding future mistakes. In order to fortify the learning the student knowledge will be tested. The language learner will get quizzes about the meaning of learned polysemous words in different contexts (Location, time).

2 DIFFICULTIES AND PROBLEMS OF FALSE FRIENDS LEARNING

From the pedagogical perspective, intrinsic and extrinsic factors determine the degree of false friends' difficulty for language (Beltran, 2006). The intrinsic factors contributing to a higher level of difficulty of false friends learning are:

IF1: The confusing nature of false friends, and particularly the partial false friends. Some false friends have always a deceptive meaning, whereas some others have deceptive meanings in certain contexts only. This creates uncertainty for students as they could fail to recognize in which contexts the word is a false friend, and in which contexts it is not.

IF2: Semantic fields may overlap. Words can have different meanings in both languages, but belong to the same semantic field (e.g.: Japanese: フィルム. firumu means camera film roll).

IF3: Because of the large number of true friends, students have a tendency to overgeneralize the words that they come across.

The extrinsic factors contributing to the complexity of false friends learning are:

EF1: Language learners are usually encouraged to take advantage of true cognates without being warned of the existence of false friends. This could lead to frustration for the language learner when they notice the actual complexity of the cognates.

EF2: Oversimplification of dictionaries by lexicographers where translations sometimes lack of nuances and contexts.

Depending on the type of false friend, the previous factors influence differently the difficulty in learning them. In order to improve false friends teaching, it is important to understand which type of factors influence false friends learning. Table 1 shows which factors affect which type of false friends learning as well as the overall learning difficulty. X is displayed when a type of false friends learning is not affected by a factor. O is displayed when a type of false friend is affected by a factor.

Table 1: Factors affecting False Friends learning and learning difficulty.

	F2: False Friends	F3: Partial False Friends N∩T	F4: Partial False Friends T⊂N	F5: Partial False Friends N⊂T
IF1	X	O	O	X
IF2	O	X	X	X
IF3	O	O	O	X
EF1	O	O	O	O

EF2	X	O	O	O
Difficulty	Medium	High	High	Low

False friends are rarely incorporated into language classes despite the difficulties faced by language students when dealing with them. When they are pointed out by the teachers, the words' nuances are often over-simplified and downgrade the accuracy and assimilation of the meaning of the word (Hayward, 1984). In the case of partial false friends, this lack of accuracy can lead to two different situations:

- a loss of some meanings of the word in the target language.
- an addition of some meanings to the word in the target language by projecting the meaning of the word in the native language to the word in the target language.

Figure 2 shows the situations where the oversimplification leads to the loss or addition of meanings.

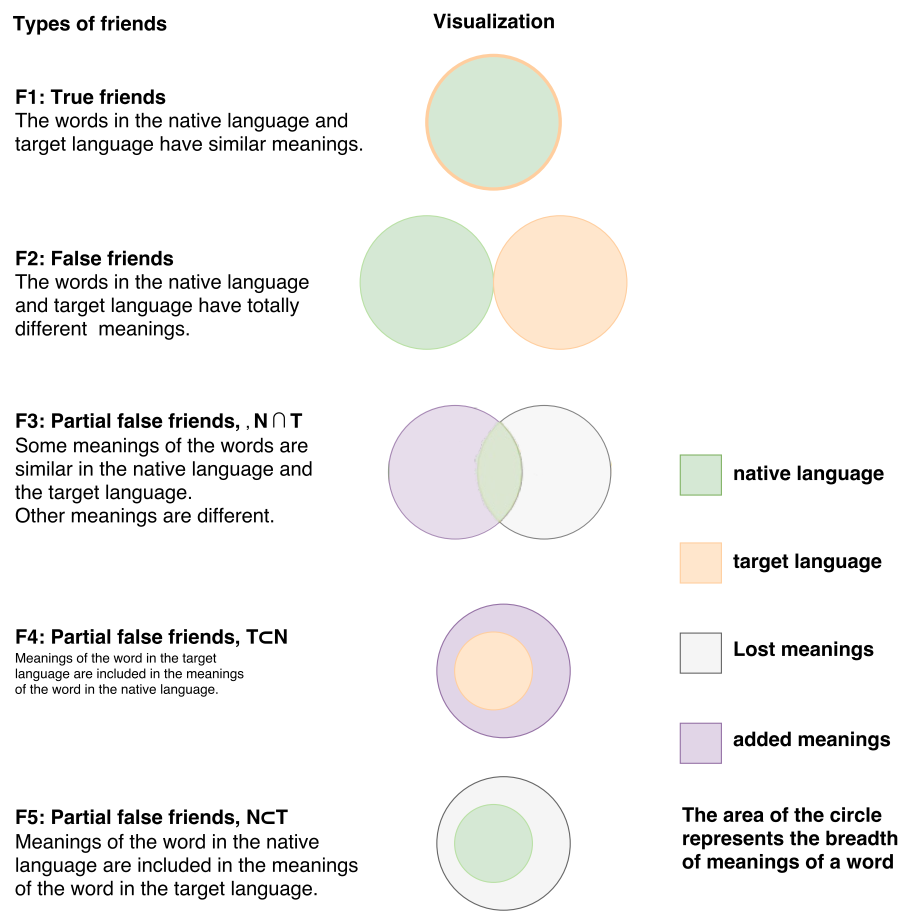


Figure 2: Loss or projection of meanings due to oversimplified teaching of false friends, depending on the type of false friend.

In order to avoid this kind of mistakes, false friends need to be pointed out by the teachers and the correct translations should be given (August, 2005). Moreover, the meanings should not be presented as one bloc to the student but should be put in a context and restricted to a particular area (Hayward, 1984).

However, in a self learning environment, teachers have less or no control over the learned content. With the rise of smart phones, the use of mobile devices in language learning is a growing trend (Godwin, 2016). Student have more control over their learning pace (Benson, 2005) and the vocabulary they want to learn, but don't have opportunities to communicate which meaning of the word they are looking for.

This paper proposes to take advantage of student past learned words to understand the particular meaning queried by the learner and provide them with the correct translation in their intended context. Moreover, we aim at minimizing the effect of the intrinsic and extrinsic factors affecting false friends learning by showing the learner the different meanings and nuances of the words. The assimilation of the student will be then tested in the context of usage of the word in order to fortify their knowledge.

3 METHOD

3.1 SCROLL system

During this study we use records from the SCROLL System (System for Capturing and Reminding Of Learning Log). Scroll is a digital record of what language students have learned in daily life. It allows the learners to log the new words or sentences they learned along with photos, audios, videos and location (Ogata, 2011). SCROLL captures what learners are learning as well as its contextual data. The users are then reminded of what they learned in the right place and the right time. Moreover, students receive personalized quizzes to fortify the learning. Figure 3 is a screenshot from the SCROLL system that shows a log inserted by a student for the word *Karaoke*. The student appended a picture and a location when creating the log. A Japanese translation of the word *Karaoke* is automatically provided to the student, and the time is automatically registered.

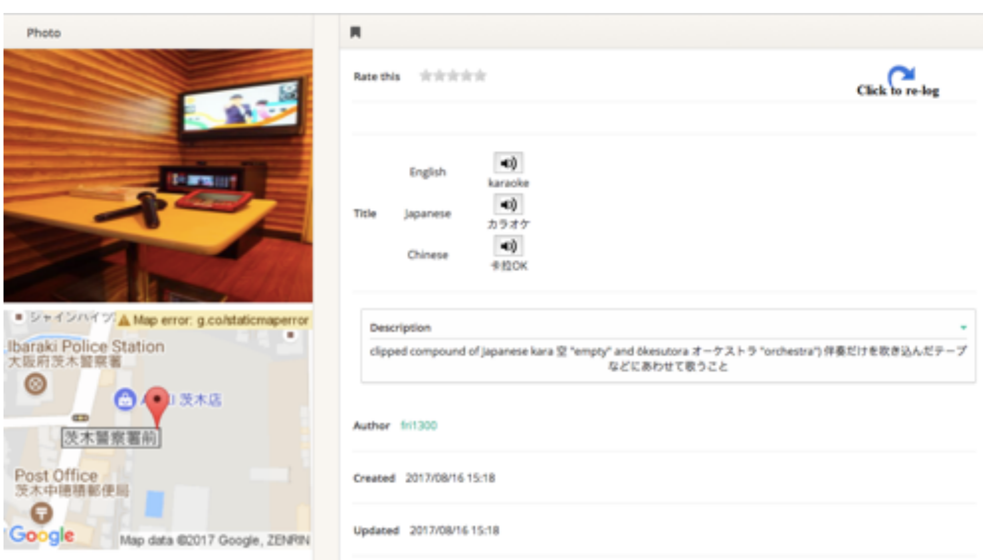


Figure 3: Screenshot from the SCROLL system showing a log inserted by a student.

The logs we will be using include meta-data such as:

Knowledge: words that students have learned in the past
User: author identification
Place: place where the learning happened (cinema, restaurant)
Time: time when the learning took place

Currently SCROLL has 1648 users and contains 24355 logs. The system is used mainly by students learning Japanese.

3.2 Contextual false friends learning

When using the SCROLL system, Japanese language learners insert logs containing a word in English and learn its Japanese translation. However, if the word is a false friend, students get a translation that does not usually reflect the context, the different meanings and the nuances of the word. In order to provide learners with the right translation in the right context, we have to understand their intended meaning. We propose to analyze the past activity of students to extract the meaning they are looking for. We assume that the past activity of a student gives a context that can be used to extract the meaning intended by the student when looking up a polysemous word. The following meta-data are used to predict the meaning the student is looking for:

Past knowledge (words): to measure the similarity between the current word the student is looking up and past words the student looked up in the system.
Time: to measure the similarity between the current word the student is looking up and the set of words the student looked up at the same period of time.
Location: to measure the similarity between the word the student is looking up and the set of words the student looked up at the same location.

We propose to measure the similarity using Jiang similarity measure. Jiang similarity measure based information content of each concept in WordNet. It assumes that each concept includes information in WordNet and the more common information two concepts share, the more similar the two concepts are (Meng, 2013). Previous studies where semantic distance measures were compared experimentally found that Jiang's measure gave the best results overall (Budanitsky, 2001).

Figure 4 shows an example. The student looked up the word *driver*. However, the word driver has different meanings and different translations in Japanese for each of those meaning:

Person who drives a vehicle: ドライバ (doraiba)
Train driver: 運転士 (untenshi)
Mass driver : マスドライバー (masu doraiba)
Computing : 仮想デバイスドライバ (kasoo debaisu doraiba)
Golf club: ドライバー (doraiba)
Screwdriver (British English):ドライバー (doraiba)

Figure 4: Screenshot from the SCROLL system showing the student past knowledge of a student looking for the meaning of the word driver

In order to understand which meaning the student is looking for, we look at the previous knowledge, the words looked up at the same period of time than the word *driver* and the words looked up at the same location of the word *driver*. In this case, the word looked up just before the word *driver* is *stepladder*. The semantic distance between the words *srewdriver* and *stepladder* is smaller than the semantic distance between the word *stepladder* and each of the words: *car driver*, *train driver*, *computer driver*, *mass driver* and *golf club*. We can conclude that the meaning the student is looking for is *screwdriver*, and that the translation that should be given is ドライバー (doraiba).

3.3 Preventive false friends learning

As stated previously, students face intrinsic and extrinsic factors that make false friend learning more difficult. We propose to tackle the factors by providing preventive false friends learning. Upon the encounter of a new false friend, the student will receive a warning that shows a comprehensive list of translations in different contexts. The displayed warning is different for each type of false friends as shown in table 2.

Table 2: Warning for different type of false friends..

False Friend Type	Type of warning received
F2: False Friends	Warning stating that the word is a false friend.
F3: Partial False Friends N∩T	Warning that: • states that the word is a partial false friend.

	• lists the different translations for the different meanings of the word. • lists the cognates of the word in the target language that have meanings that are non-existent in the native language.
F4: Partial False Friends T⊂N	Warning that: • states that the word is a partial false friend. • item lists the different translations for the different meanings of the word.
F5: Partial False Friends N⊂T	Warning that: • states that the word is a partial false friend. • lists the cognates of the word in the target language that have meanings that are non-existent in the native language.

The expected consequences of displaying warnings are:

Consequence1: Awareness that the words are used differently in different contexts: The warning states the false friend type.

Consequence2: Knowledge that avoids over-generalization (addition of meanings) : The warning lists the different translations in different contexts.

Consequence3: Knowledge that avoids over-simplification (loss of meanings): The warning lists the different translations in different contexts.

Table 3 shows how the different consequences reduce the effects of the intrinsic and extrinsic factors depending on the type of false friends.

Table 3: Consequences affecting the reduction of intrinsic and extrinsic factors depending on the type of false friends.

	F2: False Friends	F3: Partial False Friends N∩T	F4: Partial False Friends T⊂N	F5: Partial False Friends N⊂T
IF1	X	Consequence1	Consequence1	X
IF2	Consequence1	X	X	X
IF3	Consequence1	Consequence1	Consequence1	X
EF1	Consequence1	Consequence1	Consequence1	Consequence1
EF2	X	Consequence2, Consequence3	Consequence2, Consequence3	O

In the previous example, the student looked up the word driver. After predicting the intended meaning of the student, the student will be provided with the word ドライバー (doraiba) as a translation. However, *driver* and ドライバー are partial false friends of type 4 (F4). In this case, the student might use the word ドライバー (doraiba) as they would have used the word driver in English to express *Train driver*, *mass driver* or *driver in computing*. To avoid this mistake, we propose to display a warning to the student showing the different translations of the word driver, depending on the meaning (as shown in Figure 5). The warning gives the student awareness about the complexity of the word, and consciousness that the translation provided by the system can be used in particular contexts only.

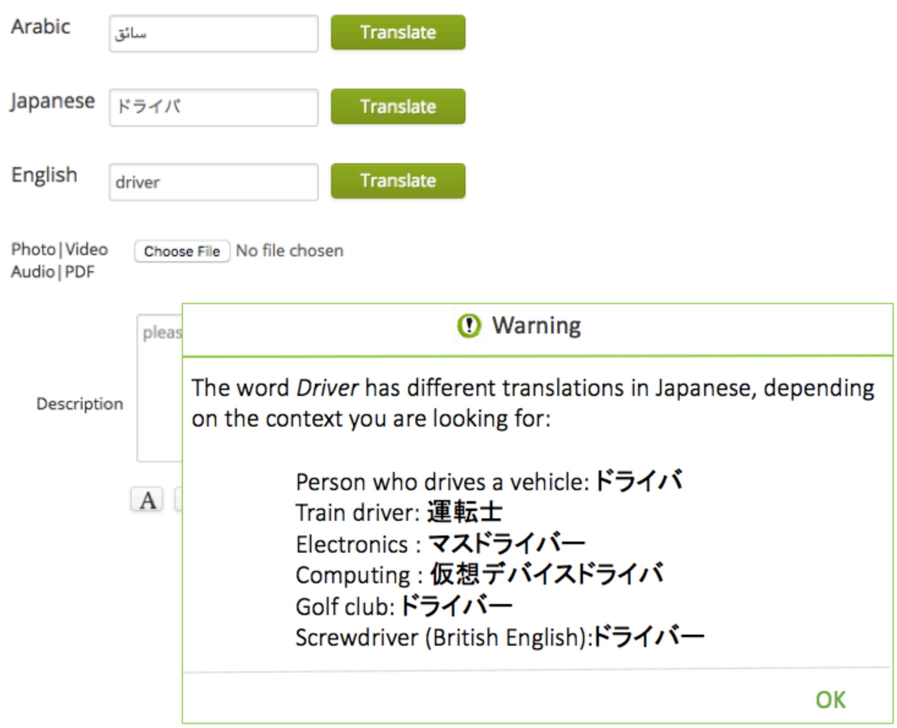


Figure 5: Warning displaying different meanings of the word depending on the context

3.4 Quizzes across contexts

In order to determine the effect of contextual translations and warnings on false friends learning, students will be given quizzes to test their acquired knowledge. SCROLL system offers the opportunity to give quizzes to students depending on their location. When the students will be present at a location related to one of the meanings of the word a quiz will appear asking the student about the translation of the word in this particular context as shown in figure 6. The quizzes will be given to the group of students that received the contextual translation as well as the warnings. Another group of students that didn't get the contextual translation and the warnings will play the role of a control group and be subjected to quizzes as well. The results will be compared to identify the effect of the previous method on false friends learning.

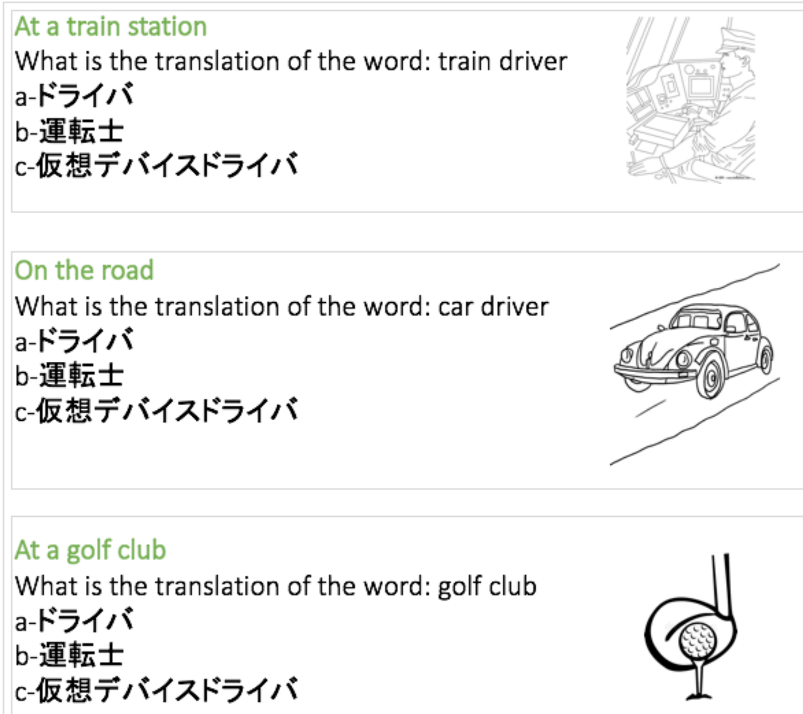


Figure 6: Quizzes displayed to the students depending on their location

4 DISCUSSION

We proposed a method to prevent false friends mistakes. The method is applied in the context of mobile learning. The main features of mobile learning are accessibility, immediacy, interactivity and situating of instructional activities (Ogata, 2004), benefit students during the learning process. However, mobile learning does not allow language students to share their intended meaning while looking up a polysemous word. The first part of our work consists of predicting the meaning intended by the student when they are looking up a polysemous word. In order to do so, the proposed method uses records from the SCROLL system (System for Capturing and Reminding Of Learning Log) to analyze the previous activity of students. We assumed that the students' past learned words, current location and time can be used to predict the meaning intended by the student when looking up a polysemous word. The identification of the intended meaning is then used to provide the student with the appropriate translation, based on the intended meaning. The second part proposes to display warnings and quizzes to the students. The warnings explain the meaning of the word and provide the student with different translations in the different contexts. The quizzes aim at fortifying the knowledge of the students, possibly improving the learning process.

This method puts into application the theoretical pedagogical approach of false friends learning. Future work will evaluate the accuracy of the prediction of the students' intended meaning. We will also evaluate the effects of the warnings by comparing the learning performance of the students before and after receiving them. The impact of the quizzes will be evaluated by comparing the false friends recall rate of the student before and after the display of the quizzes.

5 CONCLUSION

This paper proposes a method to diagnose and prevent false friends mistakes based on students' past learning activity. The proposed method uses records from the SCROLL system (System for Capturing and Reminding Of Learning Log) to analyze the previous activity of students. We assumed that the students' past learned words, current location and time can be used to predict the meaning intended by the student when looking up a polysemous word. The identification of the intended meaning is then used to provide the student with the appropriate translation, warnings and quizzes, possibly improving the learning process and avoiding future mistakes.

Acknowledgments

This research was also partially supported by Grant-in-Aid for Scientific Research (S) from Japan Society for the Promotion of Science (24220002, 2012-2016) and from Monbukagakusho (16H06304, 2016-2021).

REFERENCES

August, D., Carlo, M., Dressler, C., and Snow, C. (2005). The critical role of vocabulary development for english language learners. Learning Disabilities Research & Practice, 20(1):50–57.

Beltrán, R. C. (2006). Towards a typological classification of false friends (spanish-english). Resla, 19:29–39.

Benson, P. (2005). Autonomy and information technology in the educational discourse of the information. Information technology and innovation in language education, 1:173.

Budanitsky, A. and Hirst, G. (2001). Semantic distance in wordnet: An experimental, application-oriented evaluation of five measures. In Workshop on WordNet and other lexical resources, volume 2, pages 2–2.

Godwin-Jones, R. (2016). Looking back and ahead: 20 years of technologies for language learning.

Hayward, T. and Moulin, A. (1984). False friends invigorated. In Lexeter'83 Proceedings, pages 190–198.

Kay, G. (1995). English loanwords in japanese. World Englishes, 14(1):67–76.

Meng, L., Huang, R., and Gu, J. (2013). A review of semantic similarity measures in wordnet. International Journal of Hybrid Information Technology, 6(1):1–12.

Nation, P. (2003). The role of the first language in foreign language learning. Asian EFL Journal, 5(2):1–8.

Ogata, H., Li, M., Hou, B., Uosaki, N., El-Bishouty, M. M., and Yano, Y. (2011). Scroll: Supporting to share and reuse ubiquitous learning log in the context of language learning. Research & Practice in Technology Enhanced Learning, 6(2).

Ogata, H. and Yano, Y. (2004). Knowledge awareness for a computerassisted language learning using handhelds. International Journal of Continuing Engineering Education and Life Long Learning, 14(4-5):435–449.

Combining Multimodal Learning Analytics with Backward Design to Assess Learning

Melanie E. Peffer, Ph.D.
University of Northern Colorado
melanie.peffer@unco.edu

ABSTRACT: In this position paper, I describe a potential avenue for leveraging multimodal learning analytics research to produce evidence about how learning analytics improves learning. Recently, several members of the learning analytics community have called for an increased focus on the learning side of learning analytics, particularly in generating an evidence base. I argue here that one method for better understanding learning via analytics is to utilize a backward design approach. In backward design, an instructor begins with a specific objective and assessment and designs the pedagogical approach to meet those objectives. I extend this practice to learning analytics and suggest learning analytics design also take a backward design approach: how do we design learning analytics to fit a specific learning context and give insight into whether or not the learning objectives were achieved? By focusing specifically on the learning objective in context, this approach may advance the field by generating specific evidence for how multimodal learning analytics can be designed to assess real-time learning, rather than trying to fit existing learning analytics to the learning objective. This may lead to actionable research that could help communicate information about learning both to the student and the teacher.

Keywords: backward design, assessment, evidence, biology education

1 INTRODUCTION

Clickers, e-textbooks, adaptive reading assignments, and learning management systems are examples of technologies used in the higher-education classroom that provide potential avenues for capture of multimodal student learning data. Much of this data, such as clickstream or time on task, is easily collected and may be low-hanging fruit for understanding learning. However, what are so-called "analytics of convenience" really telling us about learning? Some have cautioned drawing rigid conclusions on findings generated solely by learning analytics to avoid too much inference of what an individual's behaviors mean (Siemens, 2015). Others have called for focusing on the *learning* side of learning analytics, rather than the easily-capturable analytics (Hackbarth, 2017). Furthermore, recent work in the learning analytics field has called for an increased focus by the community on generating evidence that learning analytics actually improves learning and pedagogy (Ferguson & Clow, 2017). In the case of educational technology, the technology notoriously comes first, causing educators to design around the technology instead of vice versa (Laurillard, 2012). How can we leverage technology to meet the specific goals of educators, instead of forcing educators to adapt to the technology? In this position paper, I propose that we can combine multimodal analytics with the principles of backward design to create deductive analytics targeting explicit research questions or learning phenomena in specific contexts. By focusing on a specific aspect on learning and intentional design of analytics to meet those goals, this may lead to more evidence for how learning analytics can improve teaching and learning in practice.

2 BACKWARD DESIGN

2.1 Backward Design and Learning

“Backward design” is a term coined by Grant Wiggins and Jay McTighe (2005) to describe a pedagogical approach where educators begin first with the desired learning outcome or result and then design the methods, materials, activities, and assessments to reach the desired learning outcome. Backward design has three distinct phases: (1) Identification of desired results, or determining what students should understand or be able to do after the unit/semester has passed; (2) Deciding what evidence, such as performance on an assessment, will demonstrate that the student achieved the desired outcome; (3) Designing appropriate instructional activities to fit the learning objectives and the method of assessment (Wiggins and McTighe, 2005). An instructor will not necessarily pass through each of these stages in order, but may cycle between them as learning activities are developed (Whitehouse, 2014). Said otherwise, in backward design the focus is on the ultimate learning goal and how that learning will be assessed instead of simply what topics need to be covered in a course, as dictated by tradition or a textbook (Wiggins & McTighe, 2005). For example, in the context of my non-majors biology course, one of my learning objectives is for students to be able to relate authentic science practices and nature of science understanding to course topics. To achieve this goal, I have specific assessments (a group project, exam items) and methods of achieving those objectives, such as completing case studies in class. My approach is backward because I started with my learning objective in mind, not with a particular project, activity, or preferred textbook. The key benefits of backward design over traditional design is students are more likely to be “hands on, minds on” rather than engaging in habitual or entertaining tasks that may not necessarily contribute to student learning.

2.2 Comparing Backward Design to Learning Design

Learning design is defined as using design knowledge when developing a learning experience, including full courses or individual lessons (Koper, 2005). Good learning experiences have good design at their base, and this design is generalizable to other learning experiences (Koper, 2005). Backward design does not necessarily have any underlying design that is generalizable to other learning experiences. If two learning experiences have similar objectives, it may be possible that one can generalize to the other. One could consider backward design as a facet of overall good learning design. When applying learning analytics to design, one method of applying useable pedagogical feedback is to design analytics to capture the learning process, or certain checkpoints to monitor student progress (Lockyer, Heathcote, & Dawson, 2013). However, this application of learning analytics to understanding pedagogy relies on using existing metrics, such as viewing student downloads from a learning management system to monitor student progress in a course or using social-network analysis to see how students complete a task (Lockyer, Heathcote & Dawson, 2013). Using a backward design approach, not only is the learning environment designed around certain objectives, but the learning analytics are intentionally designed as well around those objectives.

2.3 Backward Design and Multimodal Learning Analytics

Using a backward design approach to design of multimodal learning analytics, researchers would start with a theory-driven research question or learning phenomenon and then choose or design analytics to match the question at hand. Although analytics of convenience or extant technologies may be useful, in the context described here, their existence is considered secondary to the educational objective. In this way, we are considering “what education needs from technology” (Laurillard, 2012, p. 8) rather than what technology is available for education and research.

Use of a backward design paradigm with multimodal learning analytics parallels design-based research in that both involve the researcher working to design materials according to the specific context of interest (Barab & Squire, 2004). Multimodal analytics are of particular use since learning occurs in both digital and physical spaces, and allows a more robust method for application of backward design when choosing and implementing learning analytics.

2.4 Examples of Learning Analytics Work

In my work, we recently examined the language used by experts and novices as they engaged in simulated authentic science inquiry (Peffer and Kyle, 2017). Experts and novices differed in their expertise in authentic science practices, and we used analytics to determine which verbs were used more frequently by experts or novices. The use of expert-like hedging language is one of many sophisticated practices that my current work is pedagogically targeting. Another example of backward design in analytics is the work of Quigley, Ostwald, and Sumner (2017) which examined the modeling practices of high school students using EcoSurvey, a tool used to model ecological systems. Using modeling theory as a guide, the authors designed the analytics to capture important sequences used by the students and detect differences between teachers. Their work may provide insights in how teachers can receive personalized feedback on their instruction to promote their professional development. This is in contrast to studies such as Samson, Czarnik, and Gross (2017) or Park, Denaro, Rodriguez, Smyth, and Warschauer (2017) where easily capturable digital behaviors, such as clickstream data, were used to examine student performance. The analytics were not customized, such as in the backward design approaches used by Peffer and Kyle (2017) or Quigley, Ostwald, and Sumner (2017).

3 APPLYING BACKWARD DESIGN TO MULTIMODAL LEARNING ANALYTICS

Since assessment is a key component of backward design, using multimodal learning analytics as assessments embedded in a backward design paradigm is logical and could provide many useful insights about learning. Within the context of today's classroom, which coexists in both digital and physical spaces, using multimodal analytics could be particularly advantageous in backward design. The key benefit to using backward design in a multimodal context, with many options for capturing analytics, is to be deliberate in choosing what kinds of analytics will be the most useful to examine. How do we use each space to best capture data around a learning episode? How could devices in the physical world such as biometric sensors or smart furniture be combined with analytics in the digital world such as clickstream or natural language processing? For example, in my non-majors biology course we often discuss high profile current events such as controversial genetic technologies. Say I task students to work in groups and research a polarizing topic. Each group would then present an argument to the class, citing evidence that they found. A possible research question could be how do students choose and evaluate evidence. From the digital perspective, I could examine how many different sources of information are used, for how long they are accessed, and in what order students viewed the sources. In the physical space, I could examine language between participants around the topic at hand and biometrics. For example, what kind of biological response occurs when a student looks at contradictory information? How does this relate to their interactions with their peers? What does the data taken together tell us about learning in a multimodal space? The key differentiating factor here is starting with *what I want to know* rather than *what is already available* and designing or choosing analytic techniques to suit the learning and research objectives.

4 DISCUSSION

Although the potential for learning analytics to revolutionize research and teaching in the digital era is undisputed, there is a need for deductive, theory-driven learning analytics research to advance the field and leverage these new insights into actionable research that improves student learning outcomes (Hackbarth, 2017). Furthermore, educational needs and goals should be considered when designing analytics, and not vice-versa (Laurillard, 2012). Rather than look at easily captured data or "analytics of convenience" (e.g., clickstream data, time spent logged into a Learning Management System) and correlating these behaviors with student performance in a course, the proposed application of learning analytics here follows a backward design approach where the learning analytics are designed across physical and digital space to help achieve or assess specific learning objectives. For example, Diana et al. (2017) described how a real-time dashboard could be used by an instructor to match low and high performing students. In the hypothetical example above, an instructor could use a real-time dashboard to facilitate just-in-time teaching where the instructor views each group's progress and intervenes as needed based on the information presented on the dashboard. The analytics are intentionally designed across spaces to meet the pedagogical needs of the teacher or to provide information to the student.

Using backward design and intentionality about *what* will be collected and *why* it is important to collect will fine tune efforts to better understand learning through the use of analytics. This is particularly advantageous when considering how to meet the need in the learning analytics field to generate evidence that the learning analytics field is improving student learning (Ferguson & Clow, 2017). Although important insights about learning can be obtained via easily capturable analytics, and oftentimes this is an excellent place to start, it is also important to balance these studies with the focused, backward approaches proposed here. This may also be important when considering what methods for capturing analytics across spaces are the best investments for limited resources. Is that cool new technology fun to use, or is it going to provide important information about learning? Are we choosing a modality because it is the hot new thing (and therefore may not be that useful), or because it will help us achieve a specific goal? Wiggins and McTighe (2005) refer to these activities as "hands-on without being minds on;" learning is limited to the activity, and is not long lasting. I encourage those designing studies to consider what aspect of learning they wish to understand through learning analytics and intentionally choose what kinds of multimodal analytics to utilize. This mindset will help generate the evidence needed to give credence to the field of learning analytics, and shift the focus from the analytics to the learning.

5 ACKNOWLEDGEMENTS

I would like to thank Alan Hackbarth and David Quigley for enriching conversation on this topic and helpful comments while preparing this manuscript.

REFERENCES

Peffer, M. E., & Kyle, K. (2017, March). Assessment of language in authentic science inquiry reveals putative differences in epistemology. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 138-142). ACM.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, *13*(1), 1-14.

Diana, N., Eagle, M., Stamper, J. C., Grover, S., Bienkowski, M. A., & Basu, S. (2017, March). An instructor dashboard for real-time analytics in interactive programming assignments. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference.* ACM. (pp. 272-279).

Ferguson, R., & Clow, D. (2017). Where is the evidence? A call to action for learning analytics. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM.

Hackbarth, A. (2017). Are we Losing Sight of the Treees for the Forest? A Case for Localized Longitudinal Analytics. In Shehatea, S. & Tan, J.P-L (Eds.), *Practitioner Track Proceedings of the 7th International Learning Analytics and Knowledge Conference (LAK17)* (pp. 77-82). Simon Fraser University, Vancouver, Canada (13-17 March): SoLAR. https//solarresearch.org/wp-content/uploads/2017/02/Final-LAK17-Practitioner-Track-Proceedings.pdf

Koper, R. (2005). An introduction to learning design. In *Learning Design* (pp. 3-20). Springer.

Laurillard, D. (2013) Teaching as a design science: Building Pedagogical patterns for learning and technology. Routledge.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. American Behavioral Scientist, 57(10), 1439-1459.

Park, J., Denaro, K., Rodriguez, F., Smyth, P., & Warschauer, M. (2017). Detecting changes in student behavior from clickstream data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 21-30). ACM.

Quigley, D., Ostwald, J. L., & Sumner, T. (2017). Scientific modeling: using learning analytics to examine student practices and classroom variation. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. pp. 329-338. doi:/10.1145/3027385.3027420.

Samson, P., Czarnik, A., Gross, M. (2017) Relationships Between Digital Measures of Student Engagement and Exam Scores: Is the LMS enough? In Shehatea, S. & Tan, J.P-L (Eds.), *Practitioner Track Proceedings of the 7th International Learning Analytics and Knowledge Conference (LAK17)* (pp. 77-82). Simon Fraser University, Vancouver, Canada (13-17 March): SoLAR. https//solarresearch.org/wp-content/uploads/2017/02/Final-LAK17-Practitioner-Track-Proceedings.pdf

Siemens, G. Learning Analytics of and in Mediational Processes of Collaborative Learning. 2015 Computer Supported Collaborative Learning Conference, University of Gothenburg, Gothenburg, Sweden. 10 June 2015, Symposium.

Whitehouse, M. (2014). Using a backward design approach to embed assessment in teaching. *School Science Review*, *95*(352), 99-104.

Wiggins, G., & McTighe, T. (2005). *Understanding by Design*. (Expand 2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

On the Prediction of Students' Quiz Score by Recurrent Neural Network

Fumiya Okubo[1]*, Takayoshi Yamashita[2], Atsushi Shimada[1], Yuta Taniguchi[1] and Shin'ichi Konomi[1]

[1] Learning Analytics Center, Kyushu University, Japan

[2] *Department of Computer Science, Chubu University, Japan*

*fokubo@artsci.kyushu-u.ac.jp

ABSTRACT: In this paper, we explore the factor for improving the performance of prediction of students' quiz scores by using a Recurrent Neural Network. The proposed method is applied to the log data of 2693 students in 15 courses that were conducted with following the common syllabus by 10 teachers. The experimental results show that in the case where the same teacher is not included in both training and test data, the accuracy of prediction slightly lower. We also show that at the beginning of a course, it is better to construct a prediction model including various items of learning logs, however, in the latter half, it is better to update the model by using selected information only.

Keywords: Learning log, prediction of quiz score, recurrent neural network

1 INTRODUCTION

At Kyushu University, a learning support system called the M2B system was introduced in October 2014. The M2B system consists of three subsystems, the e-learning system Moodle, the e-portfolio system Mahara, and the e-book system BookRoll, which is enable us to record various types of logs regarding learning activities. Utilizing the collected data, various investigations on learning analytics have been conducted. The details of the M2B system and our investigations are summarized in Ogata et al. (2017) and Ogata et al. (2015).

One of the most important tasks in the field of learning analytics is to find "at-risk" students who are likely to fail or drop out of class. In Baradwaj & Pal (2011), Marbouti et al. (2016), a lot of method for this purpose are intensively investigated such as Regression, Support Vector Machine, Decision Tree. It is also valuable to identify learning activities that have significant effect on obtaining a particular final grade of students (Okubo et al. (2016), You (2016)). Our research group has developed a method for predicting final grades of students by a Recurrent Neural Network (RNN) from the log data stored in the M2B system (Okubo et al. (2017b)) which is collected from courses conducted in classrooms. An RNN is a variant of a deep neural network that handles time series data, hence it is appropriate to deal with weekly learning logs of course. In Okubo et al. (2017a), using the nine types of learning logs in multiple courses following common syllabus, the accuracy of prediction by our method is confirmed.

In this paper, we explore the factor for improving the performance of prediction by RNN. The log data was collected from 15 courses following common syllabus, which were attended by all first grade students in Kyushu University in 2017. In order to observe the prediction performance in details, the sum of quiz score of the student is regressed as the output, instead of the final grade. Moreover, we confirm the accuracy of prediction from some selected learning activities. The method and result may help teachers to find at-risk students in the course and give appropriate feedback.

2 DATA COLLECTION.

2.1 Active Learner Point

Many kinds of logs of learning activities are stored in the M2B system. To analyze and visualize these data easily, we select nine major learning activities, and evaluate them for each student from 0 to 5 points for each week of a course. The vector of these nine evaluations is called the Active Learner Point (ALP). The nine selected learning activities and the method for evaluating them are summarized in Table 1. The logs of attendance, quizzes, reports, and course views are stored in Moodle. The other logs are stored as shown in Table 1. We note that the aim of this work is to predict the sum of the quiz scores from the other learning logs. Hence, for "quiz", instead of the criteria of the original ALP, we use "0, 1" to indicate whether a student had taken a quiz.

2.2 Courses

We collected learning logs regarding ALP from 15 courses of "Basics of Cybersecurity" in the spring quarter of 2017. Each first grade student of Kyushu University was assigned to one of these courses on the basis that he/she must attend the course. These courses were conducted, basically, by following the same syllabus for eight weeks. Through attending the courses, students study entire primary cybersecurity matters including basic technologies, laws and morals of cybersecurity. Ten teachers were in charge of these courses; hence, five teachers taught two classes and the other five teachers taught one class. In Table 2, the elementary information of the courses including the number of students who attended each course is summarized. The histogram of the total score of the quizzes in eight lectures for all students in the 15 courses is shown in Figure 1.

Table 1: Active Learner Point
(the criteria for "Quiz" is not used in this paper.)

Activities	5	4	3	2	1	0
Attendance	Attendance		Being late			absence
Quiz (rate of collect answer)	Above 80%	Above 60%	Above 40%	Above 20%	Above 10%	o.w.
Report	Submission		Late submission			No submission
Course views	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	o.w.
Slide views in BookRoll	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	o.w.
Markers in BookRoll	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	o.w.
Memos in BookRoll	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	o.w.
Actions in BookRoll	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	o.w.
Word count in Mahara	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	o.w.

Table 2: Course information.

Course No.	Day	Period	Teacher	# of Students
1	Monday	3	A	204
2	Friday	3	B	206
3	Monday	3	C	188
4	Friday	3	A	217
5	Monday	4	D	201
6	Thursday	5	E	171
7	Monday	4	C	178
8	Tuesday	3	F	172
9	Thursday	3	B	170
10	Tuesday	3	G	207
11	Thursday	3	H	142
12	Tuesday	3	E	135
13	Tuesday	4	F	171
14	Thursday	3	I	151
15	Tuesday	4	J	180

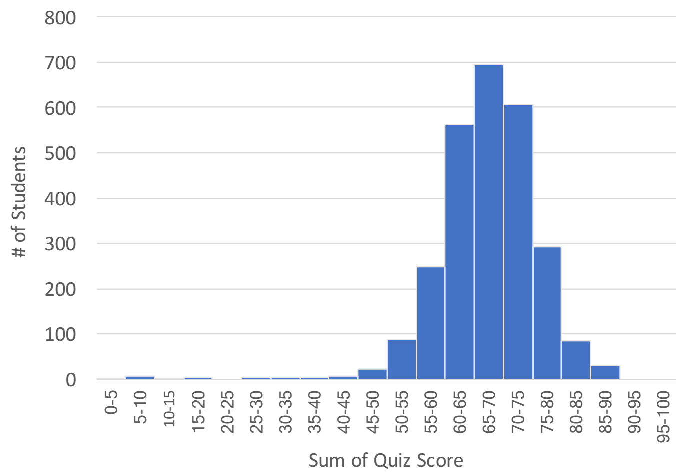
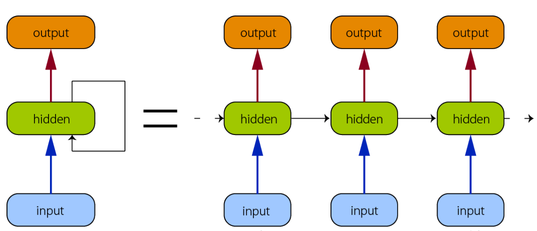


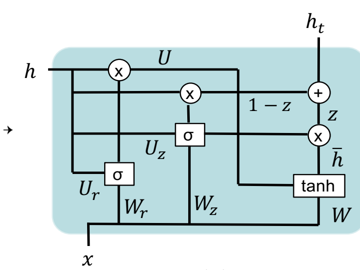
Figure 1: Histogram of the total score of the quizzes.

3 METHOD

3.1 Recurrent Neural Network

A recurrent neural network (RNN) is a variant of neural networks that handles time series data. In Figure 2 (a) shows a graphical illustration of a structure of an RNN. By inputting data to an RNN, an output value corresponding to the input value is obtained through a hidden layer. At this time, the internal information of the hidden layer based on the past data is input into an RNN, together with the information of input of the present time. Thus, it is possible to output in consideration of the past state. Figure 2 (b) shows the unfolding in the time of the computation of an RNN. Since the information of the hidden layer at time *t-1* is propagated to the same network at time *t*, an RNN theoretically can output with consideration of all the past information.





(a)

$$r = \sigma(W_r x + U_r h)$$
$$z = \sigma(W_z x + U_z h)$$
$$\bar{h} = \tanh(Wx + U(r \odot h))$$
$$h' = (1-z) \odot h + z \odot \bar{h}$$
$$\sigma(a) = \frac{1}{1+e^{-a}}$$

(b)

Figure 2: Structure of recurrent neural network. **Figure 3: Gated Recurrent Unit (GRU).**

We can select a method to construct hidden layers, such as Long Short Term Memory and Gated Recurrent Unit (GRU), depending on the way of consideration of the past information. In this paper, we deploy GRU. In Figure 3 (a), a graphical illustration of a structure of GRU is illustrated. A hidden layer of an RNN consists of *n* GRUs, where *n* is predefined. A GRU calculates intermediate values *r, z* from the input value and a value *h* of a hidden layer of the previous time. Then, the output value h_t is calculated from the value obtained by multiplying the input data by *W*, and the intermediate value *z*, and the value *h* of the hidden layer of the previous time. Through these processes, with the formulas shown in Figure 3 (b), it can be decided whether to emphasize the input data or past data, which is then reflected in the output.

3.2 Prediction of Students' Test Score

A vector of nine kinds of points for each week, that is, an ALP (introduced in Section 2.1) of a student is input into the RNN for each time. The student's quiz score from 0 to 100 is regressed as the output. Let the number of GRUs included in a hidden layer be 32. The time series data of the vectors of nine kinds of points is fed into the RNN, and in each time, the quiz score is predicted by the trained RNN. For the training of the RNN, we apply the Back Propagation Through Time (BPTT) to repeatedly update parameters of network and learn the optimal parameters.

4 EXPERIMENTS AND DISCUSSION

To evaluate the prediction performance of the proposed method, we applied cross validation for each course, that is, the data of each course was selected as the test data and the data of the rest 14 courses were treated as the training data for an RNN. For each week, a total score of quizzes were regressed for each student in the test course. We calculated the error between the predicted value and the actual quiz score. In Figure 4, the values of errors of each course for each week are summarized. Thick line represents the average of all courses. Although there is a difference depending on course, in general, we can see that the accuracy of prediction is rising as the course progresses. Looking at the average value, the error at the end of the first week is about 5.87 points, and at the end of the 8th week, it is 3.51 points. Then, in order to compare the accuracy of the case where the same teacher is included in both training data and test data with the case of not so, we extract the courses 8, 10, 11, 14, 15 by the teachers who taught in just one course and calculate the average of these courses. The result is summarized in Figure 5.

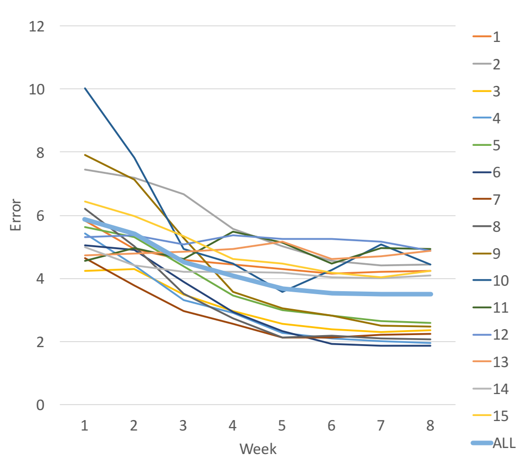


Figure 4: Difference between the predicted values and the actual data of each week.

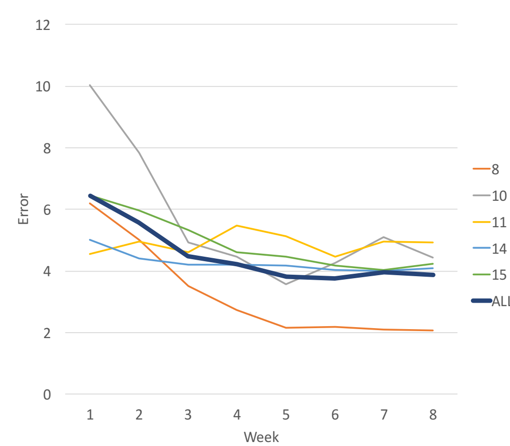


Figure 5: Difference between the predicted values and the actual data for the courses by the teachers who taught in just one course.

We also consider predicting the score of quiz from some particular items in the ALP. We selected attendance, course views and the number of actions in BookRoll that can collect from all students naturally when using the system in the course. Then, we made predictions using these three items for the courses 8, 10, 11, 14, 15. The result is summarized in Figure 6. In Figure 7, the three cases are summarized, that is, (i) the average of error of prediction using the ALP of all courses, (ii) one using the ALP of the courses 8, 10, 11, 14, 15 by the teachers who taught in just one course, and (iii) one using the data of attendance, course views and actions in BookRoll in the courses 8, 10, 11, 14, 15. Comparing the cases (i) with (ii), although the difference is not so large, it is found that the accuracy of prediction is higher in the case (i) than the case (ii) in most weeks. In the case for using the selected courses, comparing the case (iii) of using only three items in the ALP with the case (ii) of using all items in the ALP, the accuracy of (iii) is lower than (ii) in the first week, but reverses at the 4th week. At the 8th week, there is a difference of 0.67 point between (iii) and (ii). From this result, it is suggested that since there is little information at the beginning of a course, it is better to construct a prediction model including various items of learning logs, however, in the latter half, it is better to update the prediction model by continuing to collect selected important information for a long term.

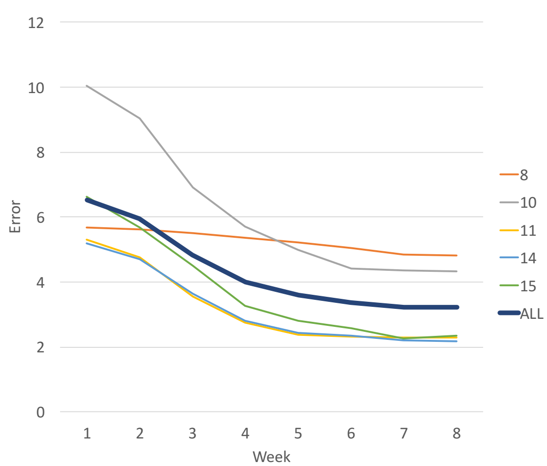


Figure 6: Difference between the predicted values by using the three items and the actual data.

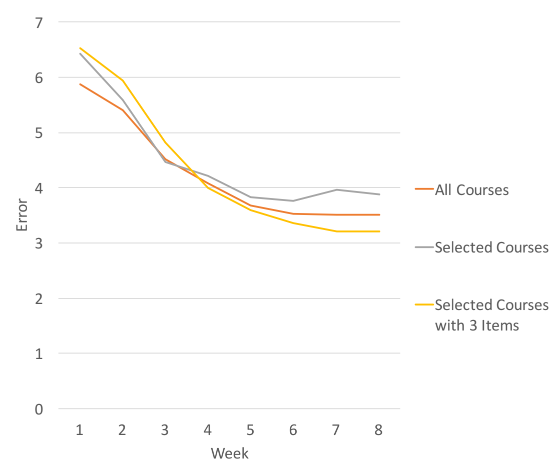


Figure 7: Difference between the predicted values the actual data for the three cases; all courses, selected courses, and selected courses with the data of attendance, course views, the number of actions.

5 CONCLUSION

In this paper, we explored the factor for improving the performance of prediction of students' quiz scores by using a Recurrent Neural Network (RNN). For this sake, the learning logs from 2693 students were collected. The nine selected learning logs stored in the M2B system are evaluated from 0 to 5 points for each student in each week of the course, and the obtained vector of these nine evaluations is called Active Learner Point (ALP). The ALPs and the total scores of quizzes are treated with input and output of RNN. The data of each course was selected as the test data and the data of the rest 14 courses were treated as the training data for RNN. Then, we calculated the error between the predicted value and the actual quiz score for each week. From the results, in the case where the same teacher is not included in both training and test data, the accuracy of prediction slightly lower. Next, we confirmed whether the accuracy of prediction do not become lower when using only selected items in ALP, that is, attendance, course views, and actions. The result suggests that it is important to investigate the method which enables us to select the optimal items of learning logs and to construct a prediction model at each time, automatically. Note that, in this paper, even though the teachers are different, training of RNN and prediction were carried out on multiple courses with the same syllabus. Hence, it is a future work to verify whether similar results can be obtained when predicting by using log data of completely different courses for training and test data.

The proposed method is shown to have high performance to predict students' quiz scores, however, to enhance students' performance using the results of prediction, the method and the timing of feedback may be an important problem to be investigated.

ACKNOWLEDGMENTS

The research results have been achieved by "Research and Development on Fundamental and Utilization Technologies for Social Big Data" (178A03), the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan.

REFERENCES

Baradwaj, B. and Pal, S. (2011). Mining Educational Data to Analyze Student's Performance, *International Journal of Advanced Computer Science and Applications*, vol. 6, 2, pp. 63-69.

Marbouti, F., H. A. Diefes-Dux and Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading, *Computer & Education*, pp. 1-15.

Ogata, H., Taniguchi, Y., Suehiro, T., Shimada, A., Oi, M., Okubo, F., Yamada, M. and Kojima, K. (2017). M2B System: A Digital Learning Platform for Traditional Classrooms in University, *Practitioner Track Proc. of LAK2017*, pp.154-161.

Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K. and Yamada, M. (2015). E-Book-based Learning Analytics in University Education, *proc. ICCE2015*, pp.401-406.

Okubo, F., Hirokawa, S., Oi, M., Shimada, A., Kojima, K. & Yamada, M. and Ogata, H. (2016). Learning Activity Features of High Performance Students, *Proc. of the 1st International Workshop on Learning Analytics Across Physical and Digital Spaces (Cross-LAK 2016)*, pp.28-33.

Okubo, F., Yamashita, T., Shimada, A. and Konomi, S. (2017a). Students' Performance Prediction Using Data of Multiple Courses by Recurrent Neural Network, *Proc. ICCE2017*, pp.439-444.

Okubo, F., Yamashita, T., Shimada, A. and Ogata, H. (2017b). A Neural Network Approach for Students' Performance Prediction, *proc. LAK2017*, pp.598-599., *proc. LAK2017*, pp.598-599.

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning, *Internet and Higher Education*, vol.29, pp.23-30.

Multimodal Transcript of Face-to-Face Group-Work Activity Around Interactive Tabletops

Xavier Ochoa, Katherine Chiluiza, Roger Granda, Gabriel Falcones, James Castells and Bruno Guamán
Escuela Superior Politécnica del Litoral, ESPOL
{xavier,kchilui,roger.granda,gabriel.falcones,james.castells,bruno.guaman}@cti.espol.edu.ec

ABSTRACT: This paper describes a multimodal system around a multi-touch tabletop to collect different data sources for group-work activities. The system collects data from various cameras, microphones and the logs of the activities performed in the multi-touch tabletop. We conducted a pilot study with 27 students in an authentic classroom to explore the feasibility of capture individual and group interactions of each participant in a collaborative database design activity. From the raw data, we extracted low-level features (e.g. tabletop action, gaze interaction, verbal intervention, emotions) and generated some visualizations as annotated transcripts of what happened in a work session. We evaluated teacher's perception about how the automated multimodal transcript could potentially support the understanding of group-work activities. Results from teachers' perceptions pointed out that the multimodal transcript could become a valuable tool to understand group-work rapport and performance.

Keywords: multimodal transcripts, collaboration, group-work visualizations

1 INTRODUCTION

With the evolution of multi-user tabletop devices, the opportunities to enhance collaboration in several contexts have been extended significantly, especially in collaborative learning contexts. Several authors have studied the effect of introducing this particular technology in collaborative sessions, reporting positive results on enhancing communication skills between participants (Kharrufa, et al. 2013; Heslop, 2015). The data-capture capabilities of tabletops in learning contexts present new opportunities to better understand the collaboration and learning processes during group-work activities in the classroom. For instance, collaboration interactions gathered from a tabletop setting could help teachers by making group-work orchestration easier (Martinez-Maldonado et al., 2011) or help students to reflect about their collaboration experience. However, using only the data produced by the tabletops provide a narrow picture of those processes because students interact through a variety of modes (speech, gaze, posture, gestures, etc.) and not all of the actions are perceived or recorded by the tabletop software (Martinez-Maldonado et al. 2017). A common approach to obtain a more holistic view of the collaboration is to complement the data captured by the tabletops with the capture and analysis from other sources such as video recordings (Al-Qaraghuli, 2013). However, the visualization of these data, especially if several communication modalities want to be captured, could become cluttered and confusing for teachers who seek to provide instant feedback to students after a group-work activity.

In this sense, an automatic multimodal transcript has proven to be an efficient and comprehensive method to represent and visualize temporal information from several sources (Bezemer & Mavers, 2011). Thus, combining multimodal collaboration features into a time-based visualization could help teachers to make sense of collaboration processes through the observation of students' actions and emotions in the group-work activity (Martinez-Maldonado et al., 2011; Tang et al., 2010). Even though some studies have added new dimensions of collaboration to the data obtained from tabletop, most of the efforts to create automated transcripts from group's interactions have been focused on one or two modalities. For example, Martinez-Maldonado et al. (2013) presented an approach to identify common patterns of collaboration by mining student logs and detected speech. In another work, Adachi et al. (2015) captured and visualized gaze and talking participation of members in a co-located conversation to provide feedback that in turn would help balancing participation. These studies serve as a baseline for our analysis; however, we want to explore the potential of generating an automated transcript by combining multiple modalities to inform teachers about groups' interactions around a tabletop. Besides, in a classroom interaction research context, the focus has been almost exclusively on teacher talk or teacher–student talk and not in student-student talk in group-work e.g. student-student rapport building in group-work (Ädel, 2011). Ultimately, we want to know if it is possible for teachers to determine more evidence about collaboration, such as group rapport from the transcripts generated.

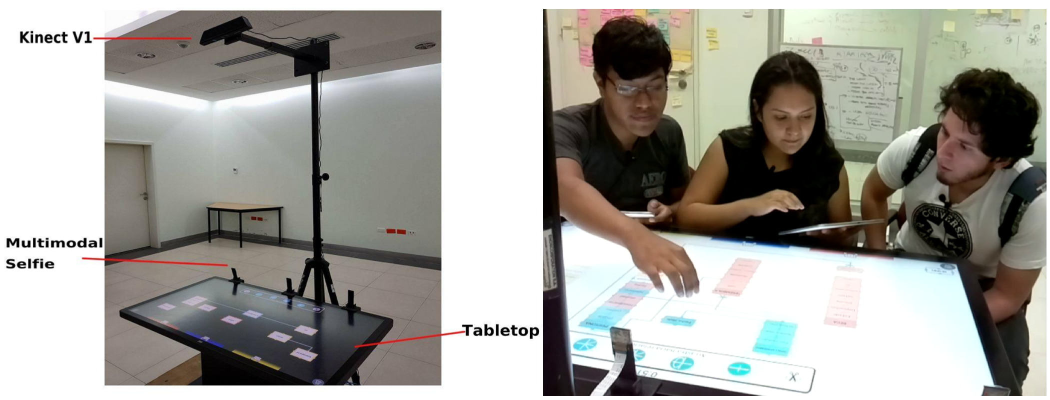


Figure 1: Components of the multimodal interactive tabletop system

2 MULTIMODAL INTERACTIVE TABLETOP SYSTEM

The system is a variation of a prototype presented by Echeverria et al. (2017), that fosters the collaborative design process. Besides the sensors used in the previous version (Kinect V1, coffee table and three tablets), this new version adds three multimodal selfies (Dominguez et al., 2015) with lapel microphones attached for capturing individual speech. The multimodal selfies are located around the tabletop to capture synchronized video and audio for each participant. Additionally, a multimodal selfie was used to capture the video of the entire group session. Figure 1 shows the components of the system and the working prototype.

The software of the system has three different applications: a tabletop application, a management web application, and a recording application. The tabletop application was designed following the design principles described in Wong-Villacres et al. (2015). It allows the participants to develop a database design through the creation and modification of several interactive objects (entities, attributes, relations). The management web application communicates with the tabletop application and with the multimodal selfies to control the execution of the session start the recordings. It also allows the instructor to view the solution developed by the students (see Echeverria, et al. 2017 for details). The recording application was deployed in each multi-modal selfie. It controls the synchronization of the recordings by the implementation of a publish/subscribe solution using the lightweight MQTT connectivity protocol[1]. The recordings obtained are further processed to automatically tag them according to the proposed audio and video features (see section 2.1). All logs and features are stored in a relational database. In addition, raw audio and video data are saved in a NAS Server, associated with a code for identifying each student.

2.1 Multimodal Transcript

The multimodal transcript combines a set of automatic features (extracted from video, audio and tabletop action logs) into a timeline where the teacher can observe the moment each interaction took place, and how the group session developed through time. The following sections present the set of features and details of the multimodal transcript.

2.1.1 Audio Features

From the speech recorded by the system, we used a Speech-to-Text recognition software, to obtain an automated transcript of the conversation among participants in the group. Then, we extracted the speech sections from the recorded individual audio of each participant, and then converted to text using Google's Cloud Speech API[2]. In this way, we obtained the conversation between the participants along with the time each verbal interaction took place. Google's API results using Spanish language are not as accurate as results obtained using English language. In spite of that, it is still useful to retrieve sentences with words related to the design problem.

2.1.2 Video Features

Mutual gaze and smiles has been considered as non-verbal indicators of rapport in previous work, (Harrigan et al., 1985). Thus, we believe that those features could be a valuable feature to be depicted in the multimodal transcript. Key points from the face of each participant recorded by the Multimodal Selfie were extracted using the OpenPose Library (Cao et al., 2016). This library retrieves the coordinates of 20 points (e.g. eyes, nose, ears, etc.). These face key points were analyzed on every frame of the recordings, and an algorithm was developed to automatically estimate the moments when a participant is looking towards to another. This evaluation was carried out by counting false positives and false negatives from all detections made by the algorithm. To evaluate the accuracy, we considered 5 different videos of 5 minute-length from the original group sessions (see section 4 for details). According to our evaluation, this feature has an error rate of 15.1%. In addition, we extracted the emotions each participant demonstrated during the activity. Video frames of individual recordings from the Multimodal Selfie were processed using the Microsoft Emotion API[3]. Thus, for every second, one frame of the participant's face is sent to the API, which returns an array of scores determining levels of happiness, anger, disgust, among others, with values from 0 to 1. To evaluate the accuracy of the emotion recognition software, videos of three students (25 min approx.) were randomly

[1] http://mqtt.org/

[2] https://cloud.google.com/speech/

[3] https://azure.microsoft.com/en-us/services/cognitive-services/emotion/

selected from all the groups that participated in a pilot study (see section 4 for details). We selected happiness as the emotion to be evaluated because it was the most common detected emotion in the recorded sessions. A human evaluated the videos by annotating if the student was happy or not for each second. Since the API returns values between 0 and 1, we selected a threshold above 0.5 to determine if the student's emotion corresponded to happiness. Our evaluation resulted in an average error rate of 1.77%.

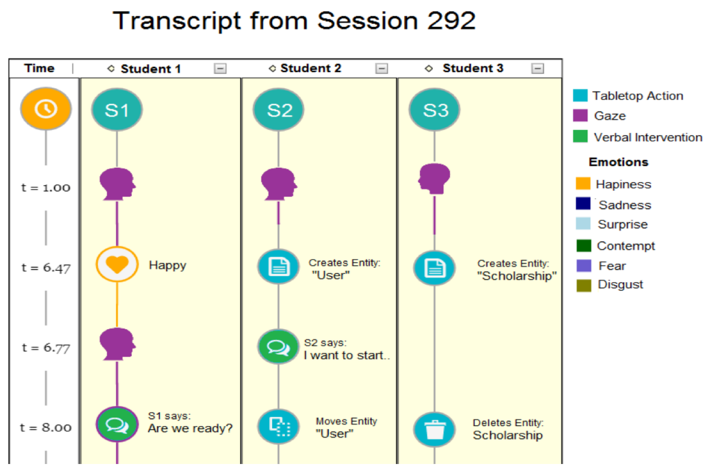


Figure 2: An excerpt of the multimodal transcript from a group

2.1.3 Interactive Tabletop Logs

All the interactions with the objects on the tabletop were recorded in a database. Each interaction is represented by the following features: type of interaction (CREATE, EDIT, DELETE), student identity, timestamp, and the type of object the student created. The solution proposed by Martínez, R. et al. (2011) was used for student differentiation while interacting with an object on the tabletop. Figure 2 shows an excerpt of a group session captured by the system. As we can see, different features of the group's interaction are represented (e.g. tabletop actions, gaze, etc.) in a vertical timeline. For instance, we can observe that in t = 1, student 1 (S1) and student 2 (S2) were looking to the right, student 3 (S3) was looking to the left, and so on.

3 PILOT STUDY

The purpose of this study was to validate with teachers the results obtained from the proposed multimodal transcript gathered from groups' sessions. The study is divided in two parts. In **a first part** of the study, twenty-four undergraduate students from a Computer Science program (20 males, 4 females, average age: 24 years), enrolled in an introductory Database Systems course and were asked to participate in a collaborative session using the proposed system. Eight groups were conformed (three students each) and grouped by affinity. Each group worked approximately 30 minutes in the design session. All of the multimodal features were recorded while the students were solving the design problem. At the end of the session, the system scored the solution proposed by the group and the teacher gave feedback to students about their performance. In addition, each student reported the rapport of their group according to their *Enjoyable Interaction* and *Personal Connection* (Frisby et al., 2010).

In the **second part** of the study, four teachers (3 males, 1 female, avg. age 31) with previous experience on teaching Database Design, were invited to participate in the evaluation of the multimodal transcript. First, teachers watched a video showing how the multimodal tabletop system works. Then, they observed the multimodal transcripts from the data gathered from two groups, corresponding to the highest and lowest rapport scores from self-reported data. For purposes of simplicity, only relevant fragments and a summary of the transcript were used in the observations. Next, the teacher answered a set of questions about the perception of *Enjoyable Interaction* and *Personal Connection* regarding each group. They assigned a score to each group for each variable using a three-point Likert Scale (Low: 1, Medium: 2. High. 3). Additionally, we included open questions about how the multimodal transcript could potentially provide support to the teacher to evaluate and recreate the group-work performance.

4 RESULTS AND DISCUSSION

As for *Enjoyable Interaction*, teachers perceived a low interaction in the group with lowest rapport, for instance they reported an average of 1.25 over 3; whereas in the group with highest rapport, they scored an average of 2.75 over 3. As for *Personal Connection*, a similar pattern was observed, the group with lowest rapport was evaluated with an average of 1.25 over 3, and the one with highest rapport scored 2.5 over 3. From these results, it seems that the multimodal transcripts were valuable for the teachers, since they mostly agreed with the rapport reported by the members of the groups. Some positive comments about the support the teachers perceived from the multimodal transcript for assessing enjoyable interaction and personal connection are presented as follows. One teacher stated*: "the combination of voice and emotions presented in the transcript give me the idea of how the students felt about the task during the session"*, another teacher said that: *"what I observed gives me evidences about the interaction among students … more specifically, the mix between actions and emotions are the evidence of such interactions"*. In addition, there were also some critical remarks. For instance, one teacher indicated that the transcript*: "did not present enough details about the emotions of the participants"*. Another teacher said that *"the emotions presented are not enough to infer the interactions that were present, I think there's the need to evidence the interrelations between the emotions of one participant with the others".*

As for the perception of teachers about how the transcript would support them to evaluate the group-work performance, all the interviewed participants answered positively to this question. Regarding the recreation of student work during the session using the multimodal transcript, three teachers answered positively and one indicated that the transcript would partially support this task. During the interviews one teacher stated that: *"I could observe whether the students were working on the task or they were debating about the task; moreover, I can observe the actions at the level of the individual. It is easy to identify who is the one who work the most or if the task was equally distributed"*. Another teacher suggested the following: *"It would be nice that the students' comment could be analyzed as well at the level of emotions"*. One teacher had a slightly reluctant reaction about the recreation of student work*: "I think it is still ambiguous what the emotions reflect in the transcript; however, the actions performed using the tabletop could help me in the recreation of the work"*.

The validation stage of this work points out to a promising research path. Teachers were mostly positive about the potential of the multimodal transcript to support group-work evaluation, beyond actions and scores. Teachers valued the fact that emotions were present in the transcript. They thought that the mix of this feature with voice and task would support the inference of interactions between the members of the groups. Nevertheless, this work is an on-going project that needs to further explore how to expand the meaning of emotions in the task, as well as, the reactions between the members of the groups after some enacted emotions.

REFERENCES

Ädel, A. (2011). Rapport building in student group work. Journal Of Pragmatics, 43(12), 2932-2947. http://dx.doi.org/10.1016/j.pragma.2011.05.007.

Adachi, H., Haruna, A., Myojin, S., & Shimada, N. (2015). ScoringTalk and watchingmeter: Utterance and gaze visualization for co-located collaboration. SIGGRAPH Asia 2015

Al-Qaraghuli, A., Zaman, H. B., Ahmad, A., & Raoof, J. (2013, November). Interaction patterns for assessment of learners in tabletop based collaborative learning environment. In *Proceedings of the 25th Australian Computer-Human Interaction Conference.* (pp. 447-450). ACM.

Bezemer, J., & Mavers, D. (2011). Multimodal transcription as academic practice: a social semiotic perspective. *International Journal of Social Research Methodology*, *14*(3), 191-206.

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2016). *Realtime multi-person 2d pose estimation using part affinity fields.* arXiv preprint arXiv:1611.08050.

Domínguez, F., Chiluiza, K., Echeverria, V., & Ochoa, X. (2015). Multimodal selfies: Designing a multimodal recording device for students in traditional classrooms. In *Proceedings of the 2015 ACM on Intl. Conf. on Multimodal Interaction* (pp. 567-574). ACM.

Echeverria, V., Falcones, G., Castells, J., Martinez-Maldonado, R., & Chiluiza, K. (2017). *Exploring on-time automated assessment in a co-located collaborative system.* Paper presented at the 4th Intl. Conf. on eDemocracy and eGovernment*, ICEDEG 2017,* 273-276.

Frisby, B. N., & Martin, M. M. (2010). Instructor–student and student–student rapport in the classroom. *Communication Education*, *59*(2), 146-164.

Harrigan, J. A., Oxman, T. E., & Rosenthal, R. (1985). Rapport expressed through nonverbal behavior. Journal of nonverbal behavior, 9(2), 95-110.

Heslop, Philip & Preston, Anne & Kharrufa, Ahmed & Balaam, Madeline & Leat, David & Olivier, Patrick. (2015). Evaluating Digital Tabletop Collaborative Writing in the Classroom.

Kharrufa, A., Balaam, M., Heslop, P., Leat, D., Dolan, P., & Olivier, P. (2013). *Tables in the wild: Lessons learned from a large-scale multi-tabletop deployment.* Conf. on Human Factors in Computing Systems, 1021-1030.

Martinez-Maldonado, R., Collins, A., Kay, J., and Yacef, K. (2011) Who did what? who said that? Collaid: an environment for capturing traces of collaborative learning at the tabletop. ACM *Intl. Conf. on Interactive Tabletops and Surfaces, ITS 2011*, pages 172-181.

Martinez-Maldonado, R., Kay, J., & Yacef, K. (2013). An automatic approach for mining patterns of collaboration around an interactive tabletop10.1007/978-3-642-39112-5-11

Martinez-Maldonado, R., Kay, J., Buckingham Shum, S. J., & Yacef, K. (2017). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. *Human-Computer Interaction.*.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2016). Detecting collaborative dynamics using mobile eye-trackers. Proceedings of International Conference of the Learning Sciences, ICLS, 1, 522-529.

Tang, A., Pahud, M., Carpendale, S., & Buxton, B. (2010). VisTACO: visualizing tabletop collaboration. In *ACM International Conference on Interactive Tabletops and Surfaces* (pp. 29-38). ACM.

Wong-Villacres, M., Ortiz, M., Echeverría, V., & Chiluiza, K. (2015). A tabletop system to promote argumentation in computer science students. In *Proceedings of the 2015 Intl. Conf. on Interactive Tabletops & Surfaces* (pp. 325-330). ACM.

Multimodal Learning Analytics' Past, Present, and, Potential Futures

Marcelo Worsley
Northwestern University
2120 Campus Drive
Evanston, IL, USA
marcelo.worsley@northwestern.edu

ABSTRACT: The first workshop on Multimodal Learning Analytics took place in 2012 at the International Conference of Multimodal Interaction in Santa Monica, California. Since then researchers have been organizing annual workshops, tutorials, webinars, conferences and data challenges. This paper examines the body of research that has emerged from these workshops, as well as in other published proceedings and journals. Within this review, a distinction is drawn between empirical and position/review/dataset papers, and looks at their rate of publication over the past several years. Additionally, key characteristics of empirical papers as well as current trends from non-empirical papers provide key insights from which to reflect on the progress of MMLA to date, and identify potential future opportunities. Specifically this review suggests that greater attention should be paid to using deep learning, developing simpler data collection tools, and finding ways to use MMLA to support accessibility.

Keywords: Machine learning, data mining, accessibility

1 INTRODUCTION

For the past several years, researchers have been conducting work in Multimodal Learning Analytics (MMLA). Worsley and Blikstein (2011) described learning analytics as "a set of multi-modal sensory inputs, that can be used to predict, understand and quantify student learning." The term MMLA was first published in 2012 at the International Conference on Multimodal Interaction (ICMI) (Scherer, Worsley, & Morency, 2012b; Worsley, 2012). Worsley, Abrahamson, Blikstein, Grover, Schneider and Tissenbaum (2016) would later elaborate on MMLA as follows:

> Multimodal learning analytics (MMLA) sits at the intersection of three ideas: multimodal teaching and learning, multimodal data, and computer-supported analysis. At its essence, MMLA utilizes and triangulates among non-traditional as well as traditional forms of data in order to characterize or model student learning in complex learning environments.

At its essence MMLA aims to leverage data from non-traditional modalities in order to study and analyze student learning in complex learning environments. In an effort to advance research within this domain, researchers have been organizing workshops and data challenges for the past five years (Morency, Oviatt, Scherer, Weibel, & Worsley, 2013; Ochoa, Worsley, Chiluiza, & Luz, 2014; Ochoa, Worsley, Weibel, & Oviatt, 2016; Scherer, Worsley, et al., 2012b; Spikol et al., 2017; Worsley, Chiluiza, Grafsgaard, & Ochoa, 2015). Additionally, a number of tutorial workshops were conducted in conjunction with the Learning Analytics Summer Institute and the International Conference of the Learning Science (Worsley et al., 2016). This paper examines the research that emerged in MMLA during this same time period. The papers included within this review are relevant papers that appeared in published conference proceedings (via CEUR and the ACM digital library), those published in the Journal of Learning Analytics and as retrieved through a google scholar. In all cases, the paper had to explicitly make reference to multimodal learning analytics. Examining past works will help ground my discussion of future opportunities for the field.

2 PAST

This review of the MMLA literature includes 82 papers. The first step in examining these papers was to determine which were empirical. This classification was based on whether or not the paper included an explicit study and analysis, versus those that present a position, a dataset or a review. 46 papers (Andrade, 2017; Blikstein, Gomes, Akiba, & Schneider, 2017; Chen, Leong, Feng, & Lee, 2014; Cukurova, Luckin, Millán, & Mavrikis, 2018; S D'Mello, Dowell, & Graesser, 2013; Davidsen & Vanderlinde, 2014; Di Mitri et al., 2017; Donnelly et al., 2016, 2017; Ez-zaouia & Lavou, 2017; Ezen-Can, Grafsgaard, Lester, & Boyer, 2015; Gomes, Yassine, Worsley, & Blikstein, 2013; Grafsgaard, 2014a; Grafsgaard, Wiggins, Vail, et al., 2014; Grafsgaard, Fulton, Boyer, Wiebe, & Lester, 2012; Grafsgaard, Wiggins, Boyer, Wiebe, & Lester, 2014; Grover et al., 2015; Hutt et al., 2017; Kory, D'Mello, & Olney, 2015; Luzardo, Guamán, Chiluiza, Castells, & Ochoa, 2014; Mills et al., 2017; Ochoa et al., 2013; Olney, Samei, Donnelly, & D 'mello, 2017; S Oviatt, Hang, Zhou, & Chen, 2015; Sharon Oviatt & Cohen, 2013, 2014; Prieto, Sharma, Dillenbourg, & Rodríguez-Triana, 2016; Raca & Dillenbourg, 2014; Scherer, Weibel, Morency, & Oviatt, 2012; Schneider, 2014; Schneider, Abu-El-Haija, Reesman, & Pea, 2013; Schneider & Blikstein, 2015; Schneider, Pao, & Pea, 2013; Schneider & Pea, 2013, 2014, 2015; Spikol, 2017; Thompson, 2013; Vail, Grafsgaard, Wiggins, Lester, & Boyer, 2014; Worsley & Blikstein, 2011a, 2011b, 2013, 2014, 2015, 2017; Worsley, Scherer, Morency, & Blikstein, 2015) were classified as empirical, while the remaining 36 papers (Andrade & Worsley, 2017; Balderas, Ruiz-Rube, Mota, Dodero, & Palomo-Duarte, 2016; Bannert, Molenar, & Azevedo, 2017; Blikstein, 2013; D'Mello, Dieterle, & Duckworth, 2017; Domínguez, Echeverría, Chiluiza, & Ochoa, 2015; Echeverria, Falcones, Castells, Granda, & Chiluiza, 2017; Eradze, Triana, Jesus, & Laanpere, 2017; Grafsgaard, 2014b; Kickmeier-Rust & Albert, 2017; Lala & Nishida, 2012; Leong, Chen, Feng, Lee, & Mulholland, 2015; Liu & Stamper, 2017; M Koutsombogera, 2014; Martinez-Maldonado et al., 2016; Martinez-Maldonado, Power, et al., 2017; Martinez-Maldonado, Echeverria, Yacef, Dos Santos, & Pechenizkiy, 2017;

Martinez-Maldonaldo et al., 2017; Merceron, 2015; Morency et al., 2013; Muñoz-Cristóbal et al., 2017; Ochoa & Worsley, 2016; Ochoa et al., 2014, 2016; S Oviatt, Cohen, & Weibel, 2013; Sharon Oviatt, 2013; Prieto, Rodríguez-Triana, Kusmin, & Laanpere, 2017; Rana et al., 2014; Rodríguez-Triana, Prieto, Holzer, & Gillet, 2017; Scherer, Worsley, & Morency, 2012a; Spikol et al., 2017; Spikol, Avramides, & Cukurova, 2016; Turker, Dalsen, Berland, & Steinkuehler, 2017; Worsley, 2012, 2017a; Worsley, Chiluiza, et al., 2015) were labelled as non-empirical. Non-empirical papers will be considered in our discussion of present work. Examining paper publication over time (Figure 1) reiterates the important role that the workshops play in advancing research in MMLA. Specifically, 2013 and 2014 were particularly productive years in terms of empirical papers as researchers were able to utilize multimodal datasets that were made publicly available. Similarly, the two MMLA workshops that were convened at Learning Analytics and Knowledge and European Conference on Technology Enhanced Learning provided two venues for researchers to present and discuss both empirical papers and position papers.

After looking at the year each paper was published, empirical papers were coded for the modalities captured, analytic techniques utilized, dependent variable, location of the study (ecological versus laboratory), whether the study was computer-mediated, the age group of the participants and whether or not the task and analysis were collaborative.

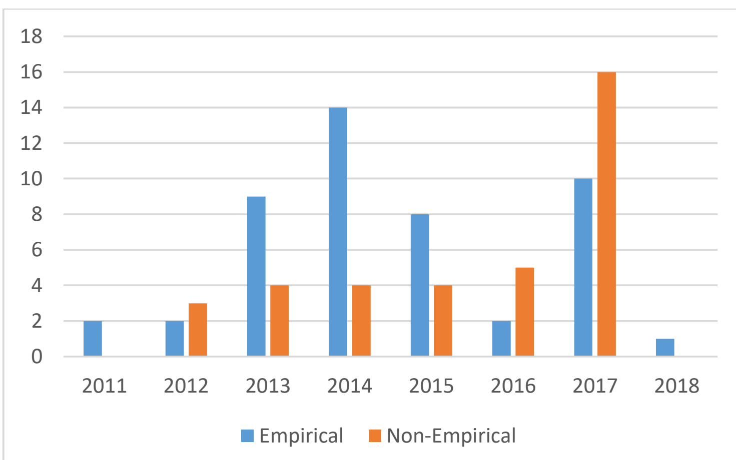


Figure 1: Empirical and non-empirical MMLA papers by year

2.1 Study Design

Study design encapsulates the nature of the task, the participants, and the location of the study.

The study design analysis begins by considering collaboration. Collaboration was coded at two levels. Specifically, the coding process recorded if the task that students completed were collaborative as well as whether or not the analysis looked at group level outcomes or individual level outcomes. Looking at the level of collaboration in the analysis will be presented later. The empirical papers included an even split between collaborative and individual tasks.

Ecological studies, i.e. those that took place outside of laboratory contexts, represented 13 of the empirical papers, while the remaining 35 were conducted in laboratory settings (Figure 2).

Additionally, 33 of the studies involved computer-mediated activities while the remaining 15 involved non-computer mediated activities.

Finally, only two of the papers worked with elementary school students. The remaining 44 working with either high school or college students (at both the undergraduate and graduate levels).

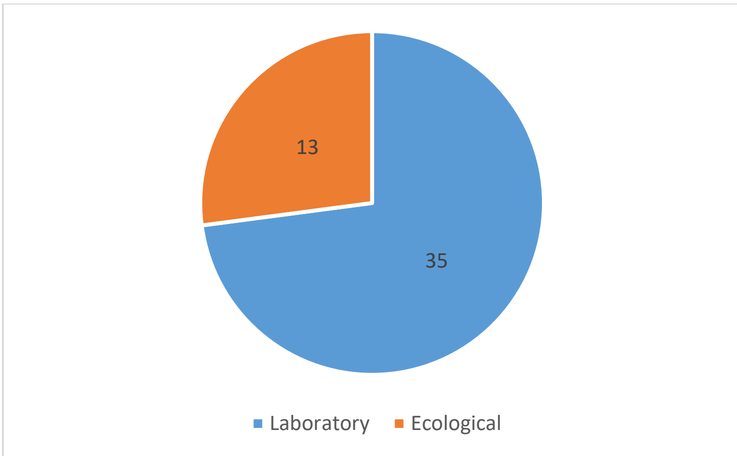


Figure 2: Ecological versus laboratory papers

2.2 Modalities Captured

Consistent with the goals of multimodal learning analytics, researchers have drawn upon a large number of modalities that include audio, video, gestures, electro-dermal activation, emotions, cognitive load and several others. The five most frequently utilized modalities are as audio, video, bio-physiology, eye tracking and digital interactions. Figure 3 includes a pie chart describing how frequently these modalities are used among the 46 empirical papers.

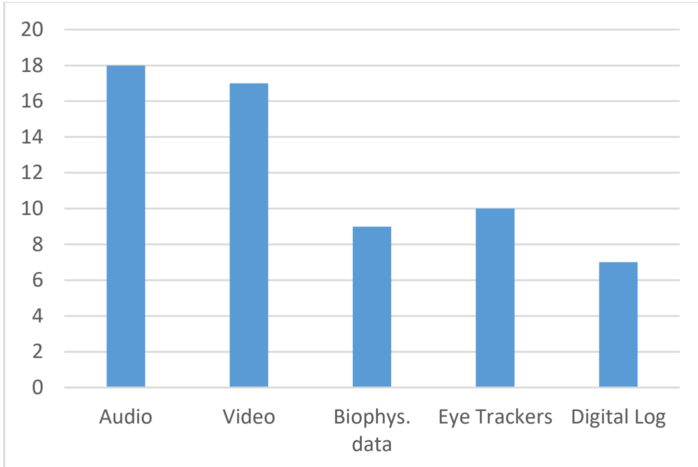


Figure 3: Number of empirical papers using most frequent modalities

Additionally, 27 of the 46 papers use at least three modalities, with the vast majority using at least two modalities. To put these modalities in context, researchers frequently used audio to analyze participant speech, and video to study body language, using both real-time and post-hoc gesture and posture tracking. Similarly, the bio-physiological measures are regularly used to study student arousal and/or affective state. Eye tracking and audio analysis are the two modalities that researchers frequently use as single modality analyses. That said, both eye tracking and audio (speech) can be easily analyzed for a variety of features (i.e., cognitive load, arousal, sentiment, entrainment).

Importantly, some of the non-empirical papers present novel data collection tools. Most notable is the Multimodal Selfies work that features synchronous two-channel audio, video and digital pen input through a Raspberry Pi.

2.3 Analysis

2.3.1 Dependent Variables

Researchers have looked at a number of constructs within their respective studies. These constructs often reflect the theoretical orientation that the researchers are following. Nonetheless, the coding of dependent variables identified several common classes of dependent variables. The constructs that emerged across multiple papers include: learning, multimodal behavior/engagement, expertise, collaboration quality, presentation quality, joint attention, affect and success (Figure 4). Understandably, the ways that researchers instantiate each of these constructs is highly variable, with some relying on human coding, while others utilize heuristics. Similarly, researchers utilize a number of different modalities to ascertain the same construct. For example, some researchers used speech signals to study affect, while others used facial expressions.

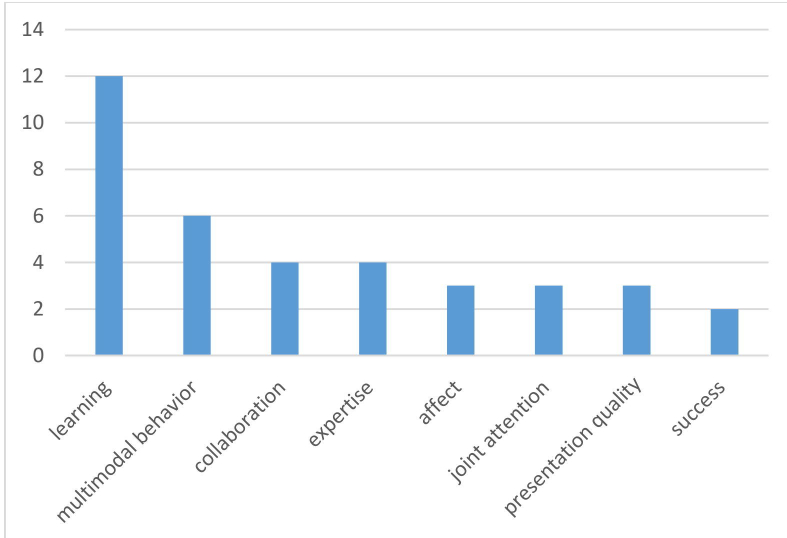


Figure 4: Number of papers using most frequent dependent variables

2.3.2 Collaboration

In considering the aforementioned learning constructs, approximately 22% looked at the group as the primary unit of analysis, 48% looked at just the individual, and 30% looked at both individuals and groups.

2.3.3 Tools and Techniques.

Some preliminary data was coded regarding the analytic techniques used to analyze the different data streams. Several of the analyses relied on custom developed scripts, though most leveraged existing code bases and/or toolkits to conduct the analyses. Examples of existing tools that researchers used include: Linguistic Inquiry Word Count (LIWC) (Tausczik & Pennebaker, 2010), FACET (previously CERT) (Littlewort et al., 2011), OpenEAR. In other cases, researchers built custom tools based on existing APIs and SDKs (e.g., Kinect for Windows, Microsoft Emotion Service API, OpenCV and the Natural Language Toolkit). Additionally, many used traditional machine learning algorithms (support vector machines (SVM), Bayesian Networks, Decision Trees, to name a few). It is also instructive to note that many of the studies used hand-annotations to seed supervised learning algorithms. However, for the sake of brevity, this review will not go into detail about each of the analytic techniques. That exposition will be left as future work.

In summary, prior work in multimodal learning analytics has been heavily geared towards studying groups of learners using a broad number of modalities completing computer-mediated tasks. The majority of the studies describe work completed in laboratory settings, and primarily includes high school and college students. Taken together, the past studies highlight the feasibility to capture and analyze multimodal data, but demonstrate this capability within a somewhat limited set of contexts and with a limited set of participant types. Turning to more present day analyses, researchers are aiming to address some of these limitations.

Note: A number of papers included within this review were based on data sets distributed in conjunction with data challenges and workshops. In particular, the 2013 and 2014 Multimodal Learning Analytics Grand Challenge Workshops were centered around the Multimodal Math Data Corpus and the Public Speaking Corpus.

3 PRESENT

Consideration of present work in MMLA is based on ideas raised in non-empirical papers from 2015 - 2017. Each document was open-coded for the central ideas that emerged. After the initial coding process like terms were grouped into categories. This section presents a summary of those categories.

3.1 Mobility

A central idea that emerged from several works (e.g., Martinez-Maldonado, Power, et al., 2017; Martinez-Maldonado, Echeverria, Yacef, Dos Santos, & Pechenizkiy, 2017; Prieto, Sharma, Dillenbourg, & Rodríguez-Triana, 2016; Rana et al., 2014; Worsley, 2012) is the ability to capture user location using mobile devices. This data allows researchers to study participants' locations in space, while also providing a relatively easy means for collecting accelerometer, video and other multimodal data streams. This data has utility for studying teacher movement within a classroom, as well as studying student-student, student-technology and student-instructor interactions.

3.2 Frameworks and Models

Another central component of contemporary MMLA is the development of frameworks and models that offer better generalizability and applicability (e.g. Andrade & Worsley, 2017; Eradze et al., 2017; Kickmeier-Rust & Albert, 2017; Liu & Stamper, 2017; Muñoz-Cristóbal et al., 2017; Prieto et al., 2017). At the same time, utilizing these frameworks can help establish norms for how data is analyzed across different contexts, and help researchers more clearly situate the objectives and orientation of their work. Finally, established frameworks and models can help with the creation of proof cases and add increased legitimacy to multimodal learning analytics research.

3.3 Data Visualization

Researchers are also looking to address challenges of data visualization, integration with existing data analysis tools, and the creation of new data analysis tools (e.g., Bannert et al., 2017; Martinez-Maldonaldo et al., 2017). While some early work and tools have been developed that help in the process (Fouse, 2011; Wagner et al., 2013), there is a significant need for new and robust tools. This heading also includes concerns related to data standardization, and the overall ease of analyzing multimodal data. Researchers have proposed utilizing existing application programming interfaces (APIs) (e.g., xAPI) (e.g., Eradze et al., 2017; Kickmeier-Rust & Albert, 2017; Prieto et al., 2017). However, as one can imagine the data standards, data visualization and data collection tools are closely connected to one another.

3.4 Human-Computer Analysis Collaboration

Another theme is opportunities to conduct research that sits at the intersection of human-computer collaboration (e.g. D'Mello et al., 2017; Spikol et al., 2016; Worsley et al., 2016; Worsley & Blikstein, 2017). Specifically, researchers are looking for ways to make the most of human inference and artificial intelligence either by bootstrapping human analysis with artificial intelligence, or periodically using human inference in the computational data analysis pipeline.

3.5 Classroom Orchestration

This category reflects current work (e.g., Martinez-Maldonaldo et al., 2017; Prieto et al., 2016) to make the output of MMLA more actionable. The action orientation can be realized through teacher and learner interfaces that participants interpret, as well as through intelligent systems that help to orchestrate the learning experience.

3.6 Cross MMLA

The current orientation of CrossMMLA reflects one of the current trends within the MMLA community. Specifically, researchers are increasingly engaged in using MMLA to study student learning across different digital and physical spaces, and in increasingly ecological, or real world, contexts. Conducting such work presents a number of novel challenges in terms of data collection, interoperability and standardization.

These categories by no means represent the entirety of current research in MMLA. However, they do touch on several of the cross-cutting ideas that are being advanced by multiple researchers within this field.

4 POTENTIAL FUTURES

Having considered the past and the present, this paper now turns to considering potential futures. Reasonably, there are several potential directions that MMLA research could take. Here, I highlight areas that may be fruitful for advancing the field, especially given the overall motivation and prior work in MMLA.

4.1 Accessibility

Absent from current work in MMLA is considerations for how these technologies can promote accessibility and inclusivity in learning. Put differently, MMLA has the potential to create novel learning experiences for people with disabilities (Worsley, 2017b), a potential that remains largely underexplored within the MMLA community. Extending MMLA to this realm is in line with many of the early motivations of MMLA. Even if the field does not yet feel prepared to leverage the available analytics to provide feedback, there is an opportunity to include more people with disabilities in the studies that we undertake.

4.2 Deep Learning

In examining the many analytic techniques currently employed in MMLA research, there appears to be significant underrepresentation of deep learning algorithms, especially given the extent to which deep learning is transforming the field of artificial intelligence. As we endeavor to stay on the cutting-edge, it will be important for the field of MMLA to find ways to leverage deep learning. In particular, several existing deep learning algorithms have the ability to easily be adapted to a specific domain on context by retraining the final layers of a deep neural network for example. In the case of computer vision, for example, the initial layers are, ostensibly, beneficial for extracting common features from an image, while the later layers are trained to handle the peculiarities of a given dataset. Researchers are also leveraging deep learning to conduct much more complex gesture tracking within large groups of people. Outside of computer vision, deep learning is also proving to be quite useful for natural language processing, both in building language models and, potentially for speaker identification. Taking advantage of these capabilities could offer a significant boost in MMLA research.

4.3 Simplifying Data Collection

In addition to thinking about accessibility and deep learning, the field is still in need of significantly simplified data collection, and analysis, tools. At present, the challenges associated with synchronously collecting multimodal data, is a significant impediment for many researchers. Too many researchers continue to rely on custom developed scripts and manual data alignment for MMLA to be tractable and accessible for those who are not already invested in this type of research.

5 CONCLUSION

This paper presented a preliminary literature review about Multimodal Learning Analytics. It used a number of criterion to identify trends in MMLA, as well as opportunities for future development. While the process for coding these papers could easily be extended to consider more details about specific data collection tools, the best approaches for multimodal triangulation or the types of analyses completed, the current literature review suggests that past and present work in multimodal learning analytics has laid a strong foundation for on-going research. Importantly, researchers have demonstrated the ability to collect multimodal data from groups of students in ecological settings, and to conduct analyses at both the individual and group levels. Moving forward the field appears to be poised to continue working in ecological settings, and to now expand towards collecting data across digital and physical spaces, while also taking advantage of the affordances of mobile technology. Furthermore, we are likely to see the development of more robust frameworks, simplified data collection and analysis tools and agreed upon standards. The field can advance this work further by taking full advantage of and contributing to deep learning. More importantly, the field would benefit for considering the ways that MMLA can positively contribute to accessibility.

REFERENCES

Andrade, A. (2017). Understanding Student Learning Trajectories Using Multimodal Learning Analytics within an Embodied- Interaction Learning Environment. *International Learning Analytics and Knowledge Conference*.

Andrade, A., & Worsley, M. (2017). A Methodological Framework for the Exploratory Analysis of Multimodal Features in Learning Activities. *Joint Proceedings of the Sixth Multimodal Learning Analytics (MMLA) Workshop and the Second Cross-LAK Workshop (MMLA-CrossLAK)*, (1828), 99–103. Retrieved from http://ceur-ws.org/Vol-1828/#paper-15

Balderas, A., Ruiz-Rube, I., Mota, J. M., Dodero, J. M., & Palomo-Duarte, M. (2016). A development environment to customize assessment through students interaction with multimodal applications. *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '16*, 1043–1048. https://doi.org/10.1145/3012430.3012644

Bannert, M., Molenar, I., & Azevedo, R. (2017). Relevance of Learning Analytics to Measure and Support Students ' Learning in Adaptive Educational Technologies. *2017 CrossLAK Workshop*. https://doi.org/10.1145/3027385.3029463

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 102. https://doi.org/10.1145/2460296.2460316

Blikstein, P., Gomes, J. S., Akiba, H. T., & Schneider, B. (2017). The Effect of Highly Scaffolded Versus General Instruction on Students' Exploratory Behavior and Arousal. *Technology, Knowledge and Learning*, *22*(1), 105–128. https://doi.org/10.1007/s10758-016-9291-y

Chen, L., Leong, C. W., Feng, G., & Lee, C. M. (2014). Using Multimodal Cues to Analyze MLA'14 Oral Presentation Quality Corpus: Presentation Delivery and Slides Quality. In *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge* (pp. 45–52). New York, NY, USA: ACM. https://doi.org/10.1145/2666633.2666640

Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Computers and Education*, *116*, 93–109. https://doi.org/10.1016/j.compedu.2017.08.007

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, Analytic, Automated (AAA) Measurement of Engagement During Learning. *Educational Psychologist*, *52*(2), 104–123. https://doi.org/10.1080/00461520.2017.1281747

D'Mello, S., Dowell, N., & Graesser, A. (2013). Unimodal and Multimodal Human Perception of

Naturalistic Non-Basic Affective States during Human-Computer Interactions. *IEEE Transactions of Affective Computing*, 1–14. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6600687

Davidsen, J., & Vanderlinde, R. (2014). Researchers and teachers learning together and from each other using video-based multimodal analysis. *British Journal of Educational Technology*, *45*(3), 451–460. https://doi.org/10.1111/bjet.12141

Di Mitri, D., Börner, D., Scheffel, M., Ternier, S., Drachsler, H., & Specht, M. (2017). Learning pulse: A machine Learning approach for predicting performance in self-regulated learning using multimodal data. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3027385.3027447

Domínguez, F., Echeverría, V., Chiluiza, K., & Ochoa, X. (2015). Multimodal Selfies : Designing a Multimodal Recording Device for Students in Traditional Classrooms. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 567–574. https://doi.org/10.1145/2818346.2830606

Donnelly, P. J., Blanchard, N., Olney, A. M., Kelly, S., Nystrand, M., & D'Mello, S. K. (2017). Words Matter: Automatic Detection of Teacher Questions in Live Classroom Discourse using Linguistics, Acoustics, and Context. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, 218–227. https://doi.org/10.1145/3027385.3027417

Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., … D'Mello, S. K. (2016). Automatic Teacher Modeling from Live Classroom Audio. *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization - UMAP '16*, 45–53. https://doi.org/10.1145/2930238.2930250

Echeverria, V., Falcones, G., Castells, J., Granda, R., & Chiluiza, K. (2017). Multimodal collaborative workgroup dataset and challenges. *CEUR Workshop Proceedings*, *1828*, 94–98. https://doi.org/10.475/123

Eradze, M., Triana, R., Jesus, M., & Laanpere, M. (2017). How to Aggregate Lesson Observation Data into Learning Analytics Datasets ? *6th Multimodal Learning Analytics Workshop (MMLA)*, 1–8.

Ez-zaouia, M., & Lavou, E. (2017). EMODA : a Tutor Oriented Multimodal and Contextual Emotional Dashboard. *Seventh International Learning Analytics & Knowledge Conference (LAK 2017)*, 429–438. https://doi.org/10.1145/3027385.3027434

Ezen-Can, A., Grafsgaard, J. F., Lester, J. C., & Boyer, K. E. (2015). Classifying student dialogue acts with multimodal learning analytics. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*, 280–289. https://doi.org/10.1145/2723576.2723588

Fouse, A. S. (2011). ChronoViz : A system for supporting navigation of time-coded data. *Chi*, 1–6. https://doi.org/10.1145/1979742.1979706

Gomes, J. S., Yassine, M., Worsley, M., & Blikstein, P. (2013). Analysing Engineering Expertise of High School Students Using Eye Tracking and Multimodal Learning Analytics. *Proceedings of the 6th International Conference on Educational Data Mining*, (Figure 1), 375–377.

Grafsgaard, J. F. (2014a). Multimodal Affect Modeling in Task-Oriented Tutorial Dialogue.

Grafsgaard, J. F. (2014b). Multimodal Analysis and Modeling of Nonverbal Behaviors During Tutoring. In *Proceedings of the 16th International Conference on Multimodal Interaction* (pp. 404–408). New York, NY, USA: ACM. https://doi.org/10.1145/2663204.2667611

Grafsgaard, J. F., Fulton, R. M., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2012). Multimodal Analysis of the Implicit Affective Channel in Computer-Mediated Textual Communication. *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, 145–152.

https://doi.org/10.1145/2388676.2388708

Grafsgaard, J. F., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2014). Predicting Learning and Affect from Multimodal Data Streams in Task-Oriented Tutorial Dialogue. In *7th International Conference on Educational Data Mining* (pp. 122–129).

Grafsgaard, J. F., Wiggins, J. B., Vail, A. K., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2014). The Additive Value of Multimodal Features for Predicting Engagement, Frustration, and Learning During Tutoring. *Proceedings of the Sixteenth ACM International Conference on Multimodal Interaction*, 42–49. https://doi.org/10.1145/2663204.2663264

Grover, S., Bienkowski, M., Tamrakar, A., Siddiquie, B., Salter, D., & Divakaran, A. (2015). Multimodal Analytics to Study Collaborative Problem Solving in Pair Programming. *LAK 2016*, 516–517. Retrieved from http://arxiv.org/abs/1505.02137

Hutt, S., Mills, C., Bosch, N., Krasich, K., Brockmole, J., & D'Mello, S. (2017). "Out of the Fr-Eye-ing Pan": Towards Gaze-Based Models of Attention During Learning with Technology in the Classroom. *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, 94–103. https://doi.org/10.1145/3079628.3079669

Kickmeier-Rust, M. D., & Albert, D. (2017). Using structural domain and learner models to link multiple data sources for learning analytics. *CEUR Workshop Proceedings*, *1828*, 82–88.

Kory, J., D'Mello, S. K., & Olney, A. M. (2015). Motion Tracker: Cost-Effective, Non-Intrusive, Fully-Automated Monitoring of Bodily Movements using Motion Silhouettes. *PLOS ONE*.

Lala, D., & Nishida, T. (2012). Joint activity theory as a framework for natural body expression in autonomous agents. *Proceedings of the 1st International Workshop on Multimodal Learning Analytics - MLA '12*, 1–8. https://doi.org/10.1145/2389268.2389270

Leong, C. W., Chen, L., Feng, G., Lee, C. M., & Mulholland, M. (2015). Utilizing depth sensors for analyzing multimodal presentations: Hardware, software and toolkits. *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction*, (3), 547–556. https://doi.org/10.1145/2818346.2830605

Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). The computer expression recognition toolbox (CERT). In *2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011* (pp. 298–305). https://doi.org/10.1109/FG.2011.5771414

Liu, R., & Stamper, J. (2017). Multimodal Data Collection and Analysis of Collaborative Learning through an Intelligent Tutoring System. *LAK Workshops '17*, 1–6.

Luzardo, G., Guamán, B., Chiluiza, K., Castells, J., & Ochoa, X. (2014). Estimation of Presentations Skills Based on Slides and Audio Features. In *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge* (pp. 37–44). New York, NY, USA: ACM. https://doi.org/10.1145/2666633.2666639

M Koutsombogera, H. P. (2014). Multimodal Analytics and its Data Ecosystem. *Proceedings of International Workshop on Roadmapping the Future of Multimodal Research*, 10–13.

Martinez-Maldonado, R., Echeverria, V., Yacef, K., Dos Santos, A. D. P., & Pechenizkiy, M. (2017). How to capitalise on mobility, proximity and motion analytics to support formal and informal education? *CEUR Workshop Proceedings*, *1828*, 39–46.

Martinez-Maldonado, R., Power, T., Hayes, C., Abdiprano, A., Vo, T., Axisa, C., & Buckingham Shum, S. (2017). Analytics Meet Patient Manikins: Challenges in an Authentic Small-group Healthcare Simulation Classroom. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 90–94. https://doi.org/10.1145/3027385.3027401

Martinez-Maldonado, R., Schneider, B., Charleer, S., Shum, S. B., Klerkx, J., & Duval, E. (2016). Interactive surfaces and learning analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 124–133. https://doi.org/10.1145/2883851.2883873

Martinez-Maldonaldo, R., Buckingham-Shum, S., Schneider, B., Charleer, S., Klerkx, J., & Duval, E. (2017). Learning Analytics for Natural User Interfaces. *Journal of Learning Analytics*, *4*(1), 24–57. https://doi.org/10.18608/jla.2017.41.4

Merceron, A. (2015). Educational data mining/learning analytics: Methods, tasks and current trends. *CEUR Workshop Proceedings*, *1443*(DeLFI), 101–109.

Mills, C., Fridman, I., Soussou, W., Waghray, D., Olney, A. M., & D'Mello, S. K. (2017). Put your thinking cap on: Detecting Cognitive Load using EEG during Learning. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, 80–89. https://doi.org/10.1145/3027385.3027431

Morency, L.-P., Oviatt, S., Scherer, S., Weibel, N., & Worsley, M. (2013). ICMI 2013 grand challenge workshop on multimodal learning analytics. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*, 373–378. https://doi.org/10.1145/2522848.2534669

Muñoz-Cristóbal, J. A., Rodríguez-Triana, M. J., Bote-Lorenzo, M. L., Villagrá-Sobrino, S. L., Asensio-Pérez, J. I., & Martínez-Monés, A. (2017). Toward multimodal analytics in ubiquitous learning environments. *CEUR Workshop Proceedings*, *1828*, 60–67.

Ochoa, X., Chiluiza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*, 583–590. https://doi.org/10.1145/2522848.2533789

Ochoa, X., & Worsley, M. (2016). Augmenting Learning Analytics with Multimodal Sensory Data. *Journal of Learning Analytics*, *3*(2), 213–219. https://doi.org/10.18608/jla.2016.32.10

Ochoa, X., Worsley, M., Chiluiza, K., & Luz, S. (2014). MLA'14: Third Multimodal Learning Analytics Workshop and Grand Challenges. *Proceedings of the 16th International Conference on Multimodal Interaction*, 531–532. https://doi.org/10.1145/2663204.2668318

Ochoa, X., Worsley, M., Weibel, N., & Oviatt, S. (2016). Multimodal learning analytics data challenges. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 498–499. https://doi.org/10.1145/2883851.2883913

Olney, A. M., Samei, B., Donnelly, P. J., & D 'mello, S. K. (2017). Assessing the Dialogic Properties of Classroom Discourse: Proportion Models for Imbalanced Classes. *Proceedings of EDM 2017*, 162–167. Retrieved from http://educationaldatamining.org/EDM2017/proc_files/papers/paper_26.pdf

Oviatt, S. (2013). Problem solving, domain expertise and learning: Ground-truth performance results for math data corpus. *Proceedings of the 15th ACM on International …*, 1–8. Retrieved from http://dl.acm.org/citation.cfm?id=2533791

Oviatt, S., & Cohen, A. (2013). Written and multimodal representations as predictors of expertise and problem-solving success in mathematics. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*, 599–606. https://doi.org/10.1145/2522848.2533793

Oviatt, S., & Cohen, A. (2014). Written Activity, Representations and Fluency as Predictors of Domain Expertise in Mathematics. *Proceedings of the 16th International Conference on Multimodal Interaction - ICMI '14*, 10–17. https://doi.org/10.1145/2663204.2663245

Oviatt, S., Cohen, A., & Weibel, N. (2013). Multimodal learning analytics: Description of math data corpus for ICMI grand challenge workshop. *2013 15th ACM International Conference on Multimodal Interaction, ICMI 2013*, 563–568. https://doi.org/10.1145/2522848.2533790

Oviatt, S., Hang, K., Zhou, J., & Chen, F. (2015). Spoken interruptions signal productive problem solving and domain expertise in mathematics. *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction*, 311–318. https://doi.org/10.1145/2818346.2820743

Prieto, L. P., Rodríguez-Triana, M. J., Kusmin, M., & Laanpere, M. (2017). Smart School Multimodal Dataset and Challenge. *CEUR Workshop Proceedings*, *1828*, 53–59. https://doi.org/10.1145/1235

Prieto, L. P., Sharma, K., Dillenbourg, P., & Rodríguez-Triana, M. J. (2016). Teaching Analytics : Towards Automatic Extraction of Orchestration Graphs Using Wearable Sensors Categories and Subject Descriptors. *International Learning Analytics and Knowledge*, 148–157. https://doi.org/10.1145/2883851.2883927

Raca, M., & Dillenbourg, P. (2014). Holistic Analysis of the Classroom. *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge - MLA '14*, *740*(Cv), 13–20. https://doi.org/10.1145/2666633.2666636

Rana, R., Reilly, J., Jurdak, R., Hu, W., Li, X., & Soar, J. (2014). Affect Sensing on Smartphone - Possibilities of Understanding Cognitive Decline in Aging Population. *Learning Analytics for and in Serious Games EC-TEL 2014 Workshop*, *8*(September), 1–7. Retrieved from http://arxiv.org/abs/1407.5910

Rodríguez-Triana, M. J., Prieto, L. P., Holzer, A., & Gillet, D. (2017). The multimodal study of blended learning using mixed sources: Dataset and challenges of the SpeakUp case. *CEUR Workshop Proceedings*, *1828*, 68–73. https://doi.org/10.1145/1235

Scherer, S., Weibel, N., Morency, L.-P., & Oviatt, S. (2012). Multimodal prediction of expertise and leadership in learning groups. *Proceedings of the 1st International Workshop on Multimodal Learning Analytics - MLA '12*, 1–8. https://doi.org/10.1145/2389268.2389269

Scherer, S., Worsley, M., & Morency, L. (2012a). 1st International Workshop on Multimodal Learning Analytics [Extended Abstract]. *International Conference on Multimodal Interaction*, 609.

Scherer, S., Worsley, M., & Morency, L.-P. (2012b). 1st international workshop on multimodal learning analytics. In *ICMI* (pp. 609–610).

Schneider, B. (2014). Toward Collaboration Sensing : Multimodal Detection of the Chameleon Effect in Collaborative Learning Settings. *Proceedings of the 7th International Conference on Educational Data Mining (EDM)*, (step 1), 435–437.

Schneider, B., Abu-El-Haija, S., Reesman, J., & Pea, R. D. (2013). Toward Collaboration Sensing: Applying Network Analysis Techniques to Collaborative Eye-tracking Data. Retrieved from http://delivery.acm.org/10.1145/2470000/2460317/p107-schneider.pdf?ip=171.67.216.22&acc=ACTIVE SERVICE&key=986B26D8D17D60C8DB53387753A2CC53&CFID=213885581&CFTOKEN=85214181&__acm__=1367805326_09890e5d339d48a83cad3ff1ffe8e77a

Schneider, B., & Blikstein, P. (2015). Unraveling Students' Interaction Around a Tangible Interface using Multimodal Learning Analytics. *Journal of Educational Data Mining*.

Schneider, B., Pao, Y., & Pea, R. (2013). Predicting Students ' Learning Outcomes Using Eye- Tracking Data. *LAK 2013*.

Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 375–397. https://doi.org/10.1007/s11412-013-9181-4

Schneider, B., & Pea, R. (2014). The Effect of Mutual Gaze Perception on Students ' Verbal Coordination. *Proceedings of the 7th International Conference on Educational Data Mining (EDM)*, (Edm), 138–144.

Schneider, B., & Pea, R. (2015). Does Seeing One Another ' s Gaze Affect Group Dialogue ? A Computational Approach LITERATURE REVIEW : WHY IS JOINT VISUAL ATTENTION AN IMPORTANT. *Journal of Learning Analytics*, *2*(2), 107–133. https://doi.org/10.18608/jla.2015.22.9

Spikol, D. (2017). Using Multimodal Learning Analytics to Identify Aspects of Collaboration in Project-Based Learning. *Cscl*, (June), 263–270. https://doi.org/10.22318/cscl2017.37

Spikol, D., Avramides, K., & Cukurova, M. (2016). Exploring the interplay between human and machine annotated multimodal learning analytics in hands-on STEM activities. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 522–523. https://doi.org/10.1145/2883851.2883920

Spikol, D., Prieto, L. P., Rodríguez-Triana, M. J., Worsley, M., Ochoa, X., & Cukurova, M. (2017). Current and future multimodal learning analytics data challenges. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, 518–519. https://doi.org/10.1145/3027385.3029437

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54.

Thompson, K. (2013). Using micro-patterns of speech to predict the correctness of answers to mathematics problems. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*, 591–598. https://doi.org/10.1145/2522848.2533792

Turker, A. G., Dalsen, J., Berland, M., & Steinkuehler, C. (2017). Challenges to Multimodal Data Set Collection in Games Based Learning Environments. *2017 CrossLAK Workshop*, 1–7.

Vail, A. K., Grafsgaard, J. F., Wiggins, J. B., Lester, J. C., & Boyer, K. E. (2014). Predicting Learning and Engagement in Tutorial Dialogue. *Proceedings of the 16th International Conference on Multimodal Interaction - ICMI '14*, 255–262. https://doi.org/10.1145/2663204.2663276

Wagner, J., Lingenfelser, F., Baur, T., Damian, I., Kistler, F., & André, E. (2013). The social signal interpretation (SSI) framework: multimodal signal processing and recognition in real-time. In *Proceedings of the 21st ACM international conference on Multimedia* (pp. 831–834).

Worsley, M. (2012). Multimodal learning analytics: enabling the future of learning through multimodal data analysis and interfaces. In *Proceedings of the 14th ACM international conference on Multimodal interaction* (pp. 353–356).

Worsley, M. (2017a). Engineering Design with Everyday Materials Multi- modal Dataset. *CEUR Workshop Proceedings*. Retrieved from http://ceur-ws.org/Vol-1828/paper-16.pdf

Worsley, M. (2017b). Multimodal Analysis. In J. Roschelle, W. Martin, J. Anh, & P. Schank (Eds.), *Cyberlearning Community Report: The State of Cyberlearning and the Future of Learning With Technology* (pp. 46–50). Menlo Park: SRI International.

Worsley, M., Abrahamson, D., Blikstein, P., Grover, S., Schneider, B., & Tissenbaum, M. (2016). Situating Multimodal Learning Analytics. *International Conference for the Learning Sciences 2016*, 1346–1349.

Worsley, M., & Blikstein, P. (2011a). Toward the Development of Learning Analytics: Student Speech as an Automatic and Natural Form of Assessment. *Annual Meeting of the American Education Research Association*.

Worsley, M., & Blikstein, P. (2011b). What's an Expert? Using Learning Analytics to Identify Emergent

Markers of Expertise through Automated Speech, Sentiment and Sketch Analysis. In *Proceedings of the Fourth Annual Conference on Educational Data Mining (EDM 2011)* (pp. 235–240). Eindhoven, Netherlands.

Worsley, M., & Blikstein, P. (2013). Towards the Development of Multimodal Action Based Assessment. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 94–101). New York, NY, USA: ACM. https://doi.org/10.1145/2460296.2460315

Worsley, M., & Blikstein, P. (2014). Deciphering the Practices and Affordances of Different Reasoning Strategies through Multimodal Learning Analytics. In *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge* (pp. 21–27). New York, NY, USA: ACM.

Worsley, M., & Blikstein, P. (2015). Leveraging Multimodal Learning Analytics to Differentiate Student Learning Strategies. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 360–367). New York, NY, USA: ACM. https://doi.org/10.1145/2723576.2723624

Worsley, M., & Blikstein, P. (2017). A Multimodal Analysis of Making. *International Journal of Artificial Intelligence in Education*.

Worsley, M., Chiluiza, K., Grafsgaard, J. F., & Ochoa, X. (2015). 2015 Multimodal Learning and Analytics Grand Challenge. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 525–529). New York, NY, USA: ACM. https://doi.org/10.1145/2818346.2829995

Worsley, M., Scherer, S., Morency, L.-P., & Blikstein, P. (2015). Exploring Behavior Representation for Learning Analytics. *Proceedings of the 2015 International Conference on Multimodal Interaction (ICMI)*, 251–258.

The Big Five: Addressing Recurrent Multimodal Learning Data Challenges

Daniele Di Mitri
Open University of The Netherlands
daniele.dimitri@ou.nl

Jan Schneider
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
schneider.jan@dipf.de

Marcus Specht
Open University of The Netherlands
marcus.specht@ou.nl

Hendrik Drachsler
Open University of The Netherlands – DIPF
hendrik.drachsler@ou.nl

ABSTRACT: The analysis of multimodal data in learning is a growing field of research, which has led to the development of different analytics solutions. However, there is no standardised approach to handle multimodal data. In this paper, we describe and outline a solution for five recurrent challenges in the analysis of multimodal data: the data collection, storing, annotation, processing and exploitation. For each of these challenges, we envision possible solutions. The prototypes for some of the proposed solutions will be discussed during the Multimodal Challenge of the fourth Learning Analytics & Knowledge Hackathon, a two-day hands-on workshop in which the authors will open up the prototypes for trials, validation and feedback.

Keywords: multimodal learning analytics, wearables, CrossMMLA, sensor-based learning

1 BACKGROUND

The Learning Analytics & Knowledge (LAK) community has acknowledged the necessity of taking into account physical and co-located learning activities as much as practice-based skills training; it is undeniable that in the classroom and at the workplace these "offline moments" still represent the bulkiest set of learning activities. Bringing these moments into account requires extending the data collection to additional data sources which go beyond the conventional ones, such as online learning systems, Massive Online Open Courses (MOOCs) platforms or student information systems. With the term *multimodal data*, we refer to the learning data sources collected "beyond user-computer interaction", i.e. those data sources collected during learning moments alternative to the classic desktop-based learning scenario. Although user-computer interaction data could still hold some relevant information, they can be complemented by additional multimodal data; these data can be classified into 1) data describing the learner's behaviour: including motoric and physiological data; 2) data regarding the learning situation, including social context, learning environment and learning activity. Most of these aspects can be monitored through wearable sensors, cameras or Internet of Things (IoT) devices. These tools can capture only what is "visible" by a generic sensor, meaning they generally do not have the ability to reason on the meaning behind the collected data. The observability line – i.e. what is visible by sensors and what not, conceptually separates multimodal data by human-driven qualitative interpretations, like expert reports or teacher assessments. The latter, that are more qualitative and human-driven, describe dimensions that the sensors cannot directly observe, such as learning outcomes, cognitive aspects or affective states.

Bridging the gap between learner's complex behavioural patterns with learning theories and other unobservable dimensions is the paramount challenge for multimodal analysis of learning (Worsley, 2014). Multimodal data can be used as historical evidence for the analysis and the description of the learning process: this field of research is called *Multimodal Learning Analytics* (Blikstein, 2013). The related literature shows the potential to apply a multimodal approach in a variety of learning settings including dialogic learning in teacher-student discourse (D'mello et al., 2015); computer-supported collaborative learning during knowledge-sharing and group discussions (Martinez-maldonado et al., 2017; Schneider & Blikstein, 2015); in practice-based and open-ended learning tasks, when understanding and executing a practical learning tasks (Ochoa et al., 2013).

The potential benefits of multimodal data are not only limited to analytics, e.g. human interpretation of dashboards or other visual metaphors. If multimodal data are reliable and correctly addressed and exploited, they can be used as the base to drive machine intelligence and achieve better personalisation and adaptation during learning. Multimodal data is expanding the horizon of the Learning Analytics community and its moving towards the intelligent tutoring and the artificial intelligence in education communities. For decades the long-term goal of these communities consisted in designing intelligent computer agents empathic to the learners which work as an *instructor in the box*, and that can implement strategies to reduce the difference between experts and student performance (Polson, Richardson, & Soloway, 1988). Multimodal data can facilitate achieving this goal, by equipping intelligent tutors with action-based recognition and reasoning, so that they can deal with open-ended learning tasks in uncontrolled environments.

2 MULTIMODAL CHALLENGES

The analysis of multimodal data in learning is a fairly new but a steadily growing field of research. As the interest tracing learning through the use of multimodal data grows, the opportunities stemming from it become more evident. As some authors have pointed out, the field of MLA faces a set of open challenges that create research gaps that need to be filled (Blikstein & Worsley, 2016). For instance, the LAK community (and its CrossMMLA interests group) still lacks a standardised approach for modelling of the evidence extracted from the learning process and producing valuable feedback with multimodal data. In contrast, multiple tailored ad-hoc solutions have been developed in related researches. A standardised approach to MMLA, in our understanding, should help researchers in setting-up their multimodal experiments by clarifying how the collection, storage, analysis and exploitation of the multimodal data takes place in a pragmatic and scalable manner that can be adopted into real-life educational settings. To contribute filling this gap, in this paper, we outline five main challenges stemming from the feedback loop empowered by multimodal data and learning analytics. For each of these challenges, we describe possible solutions or approaches. The prototyping, testing and validation of the proposed solutions, coincide with the agenda of the *multimodal challenge* proposed for the Fourth Learning Analytics Hackathon[1]. In these two-days, hands-on, pre-conference event, we will roll-out the first prototypes like the *LearningHub* or the *Visual Inspection Tool*; we will test their usability and validity and we will open them up for discussion with experts in the field.

2.1 Data collection

The first step of the journey is the data collection, that being the creation of datasets through multiple sensors and external data sources. The sensors employed are most likely to be produced by different vendors, hence to have different specifications and support. The approach used for data collection must be flexible and extensible to different sensors, it should allow the collection of data at different frequencies and formats. Strongly connected to the collection is the data synchronisation.

Proposed solution: to address this challenge, we introduce the *LearningHub*, a software prototype whose purpose is to synchronise and fuse different streams of multimodal data generated by the multiple sensor-applications. The *LearningHub*'s main role is to deal with the low-level specifications of every sensor offering a customisable interface to start and stop the capturing of a meaningful part of a learning task, i.e. moments clearly definable by atomic actions; we call this an *Action Recording*. The *LearningHub* is responsible to collect the updates for every sensor, organising and synchronising them chronologically.

2.2 Data storing

The second step is the data storing that encompasses the serialisation, storing and logic for retrieval of the action recordings. This step is crucial to organise the complexity of multimodal data which has multiple formats and big sizes.

Proposed solution: The *LearningHub* channels the data from multiple sensors and provides as output multiple JSON files, which serialise and synchronise the sensor values for each sensor application. The JSON files allow for sensors having multiple attributes with different time frequencies and formats; they work as exchange format documents and provides also the logic to facilitate the action recording for storing and later retrieval.

2.3 Data annotation

The *data annotation* challenge consists in finding a seamless and unobtrusive approach for labelling the learning process, i.e. triangulating the multimodal action recordings with the evidence (e.g. video clips) of the learning activities. The annotation step is rather crucial, as most of the time the meaning of a recording is not trivial to derive just by looking at the sensor values. The format chosen for assigning the semantics to the action recordings is also a relevant issue.

---

[1] LAK Hackathon 2018, Sydney, Australia, March 5-6, 2018, https://lakhackathon.wordpress.com/

Proposed solution: to address this challenge, we propose the *Visual Inspection Tool (VIT)*. The VIT is a web-application prototype for the retrospectively analysis and annotation of multimodal action recordings. The VIT allows to load multimodal datasets, plot them on a common time scale and triangulate them with video recordings of the learning activity. It allows to select a particular timeframe and annotate the multimodal data slice with an Experience API (xAPI) triplet, assigning an actor, a verb and an object. The VIT offers a human-computer interface which helps to deal with the complexity of multimodal data.

2.4 Data processing

The data processing steps consist in extracting and aligning the relevant attributes from the "raw" multimodal data and transforming them into a new data representation suitable for exploitation. The data processing steps depend tightly on the data exploitation which is discussed in next section. Common steps for data processing include data cleaning (e.g. handling missing values, resampling and realigning the time series), feature extraction, dimensionality reduction and normalisation. The challenging side of the data processing for multimodal data is given by the size of the multimodal datasets, the need to process them periodically and the need to process as close to real-time as possible, a relevant aspect especially in the case of immersive feedback generation.

Proposed solution: the idea is to have a Pipeline for multimodal data for learning, a cloud-based application which allows to plan and execute data processing routines (e.g. Spark jobs). These routines should query the Learning Record Store and fetch the all recent/relevant xAPI statements and load into memory all the action recordings connected to each xAPI statement. The raw action recordings will be transformed according to the set of operations specified which will output a transformed action recording which is saved and ready to be fed into a data mining algorithm.

2.5 Data exploitation

Through an analysis of the related experiments in the literature using multimodal data in learning settings, we concluded that there are different *use cases* generally used for enhancing and facilitating the learning process with multimodal data.

Proposed solution: we classify the different use cases into five *exploitation strategies:*

1. *light-weight feedback*: hardcoded rules and feedback based on heuristics of the form "if sensor value is x then y";
2. *replica*: replays of the action recordings, e.g. ghost-tracks of motoric sensors data;
3. *historical reports*: aggregated visualisations in forms of analytics dashboard, a group of data visualisations that show the historical progress of the sensor recordings in condensed form;
4. *frequent patterns*: mining of recurrent sensor values occurrences within one or multiple sensor recordings;
5. *predictions*: estimation of the human annotated labels during similar action recordings.

The strategies can be used for different purposes and applications. They differ in the level of data processing used and consequently by the methods used for data analysis; these include descriptive statistics, supervised or unsupervised machine learning. For example, *light-weight feedback* requires simple hardcoded rules; *historical reports* require visualisations that can be grouped into analytics dashboard; *frequent patterns* or *predictions* require training either machine learning models, store them into memory, and use them to estimate the value or the class of a particular target attribute. Historical reports also differ by the effort required by human experts, for example in collecting the labels or for interpreting the visualisations; in a similar way, the strategies differ by the level of machine reasoning, e.g. between those using machine learning and those which use heuristics.

3 CONCLUSIONS

In this paper, we have introduced five main challenges connected to the use of multimodal data in learning. These challenges deal with the data collection, storing, annotation, processing and exploitation and constitute important research questions for all the CrossMMLA community. Along with these challenges, we briefly explained some practical solutions. Being these ideas preliminary, we use them as agenda points and research questions to the Multimodal Challenge of the LAK Hackathon, a hands-on workshop which will take place during the eight Learning Analytics & Knowledge Conference in Sydney. We hope that pointing out these challenges can raise interest and awareness in the current research endeavours in the area of multimodal learning analytics.

REFERENCES

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 102. http://doi.org/10.1145/2460296.2460316

Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining : using computational technologies to measure complex learning tasks, *7750*(x), 220–238. http://doi.org/http://dx.doi.org/10.18608/jla.2016.32.11

D'mello, S., Olney, A., Blanchard, N., Sun, X., Ward, B., Samei, B., & Kelly, S. (2015). Multimodal Capture of Teacher-Student Interactions for Automated Dialogic Analysis in Live Classrooms. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 557–566. http://doi.org/10.1145/2818346.2830602

Martinez-maldonado, R., Power, T., Hayes, C., Abdiprano, A., Vo, T., & Shum, S. B. (2017). Analytics Meet Patient Manikins : Challenges in an Authentic Small-Group Healthcare Simulation Classroom, 1–5. http://doi.org/10.1145/3027385.3027401

Ochoa, X., Chiluiza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI '13)*, 583–590. http://doi.org/10.1145/2522848.2533789

Polson, M. C., Richardson, J. J., & Soloway, E. (1988). *Foundations of intelligent tutoring systems. Foundations of intelligent tutoring systems.* Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1988-97172-000&lang=de&site=ehost-live

Schneider, B., & Blikstein, P. (2015). Unraveling Students' Interaction Around a Tangible Interface using Multimodal Learning Analytics. *JEDM - Journal of Educational Data Mining*, *7*(3), 89–116.

Worsley, M. (2014). Multimodal learning analytics as a tool for bridging learning theory and complex learning behaviors. *3rd Multimodal Learning Analytics Workshop and Grand Challenges, MLA 2014*, 1–4. http://doi.org/10.1145/2666633.2666634

DesignLAK18: Evaluating systems and tools that link learning analytics and learning design

Linda Corrin
University of Melbourne
lcorrin@unimelb.edu.au

Nancy Law
University of Hong Kong
nlaw@hku.hk

Ulla Ringtved
University College Northern Denmark
ulr@ucn.dk

Sandra Milligan
University of Melbourne
s.milligan@unimelb.edu.au

ABSTRACT: The 3rd DesignLAK workshop focuses on evaluation of the frameworks, systems and tools that bring together learning design and learning analytics. The purpose of the proposed workshop is to bring together members of the learning analytics community to share their experiences of developing and researching frameworks, systems and tools that link these two fields together. This hands-on workshop will give participants a chance to use and explore a range of systems and tools that link learning analytics with learning design, applying use case scenarios to evaluate the strengths and weakness of the systems. Throughout the day an ongoing conversation will be held to identify opportunities and challenges in moving the field forward. The outcomes of the workshop will include evaluative feedback for system/tool developers as well as a discussion paper summarising the key insights that emerge from the workshop.

Keywords: Learning design, learning analytics, evaluation

1 BACKGROUND

This workshop focuses on evaluating the frameworks, systems and tools in the learning design and learning analytics communities that seek to bring these two fields together. The connection between learning analytics and learning design has been recognised as fundamental in enabling understanding and translation of the outcomes of analytics processes so they can be applied to support student learning. Learning design involves both the process of designing sequences of learning activities for students as well as the description of these learning activities for teachers' use (Bennett, Agostinho & Lockyer, 2017). The creation of a learning design enables a teacher and/or learning designer to articulate the pedagogical intent of the learning activities in sharable ways (Lockyer, Heathcote & Dawson, 2013).

When combined with learning analytics, learning design has the potential to "provide a semantic structure for analytics" (Mor, Ferguson & Wasson, 2015) helping teachers to link analytics outcomes with the pedagogy and design that underpinned the learning activity. Many opportunities are opened up by this connection including the ability for teachers to intervene in real-time to provide student support, the ability to make learning outcomes visible to teachers and students, and to provide evidence on which to base learning design decisions (Schmitz et al., 2017). A core challenge also exists in finding a common vocabulary that teachers and learning designers can use to represent their learning designs in order that they can be shared with others (Law et al., forthcoming).

Over recent years several frameworks that connect learning analytics and learning design have emerged (e.g. Bakharia et al., 2016; Donald et al., 2016; Schmitz et al., 2017). Each of these frameworks proposes ways that learning design can be integrated with the process or cycle of learning analytics to provide insight to stakeholders (e.g. teachers, students, administrators, etc.). Implicit in each of these models is the role of the teacher in translating the analysis into action. Yet, these frameworks all operate at a broad level and do not attempt to dictate a particular representation for learning designs or vocabulary through which they can be described. Consequently, while acknowledgement of the importance of the role of learning design in learning analytics has been established, there is still work to be done to operationalise this in meaningful and actionable ways for stakeholders.

Within the learning analytics community there have been several projects that have developed systems and tools to connect learning analytics with learning design (for example, Corrin et al., 2016; Law et al., forthcoming; Persico & Pozzi, 2015). There has also been research that has used learning analytics to explore the relationship between learning design and academic outcomes such as student satisfaction and performance (e.g. Rienties & Toetenel, 2016). All these endeavors demonstrate the potential offered by combining learning analytics and learning design in ways that can help improve learning environments. The purpose of the proposed workshop is to bring together members of the learning analytics community to share their experiences of developing and researching frameworks, systems and tools that link learning analytics with learning design. This workshop continues the conversation begun in previous DesignLAK workshops which have focused on stocktaking emergent theory and practice of learning design and feedback processes (Milligan et al., 2016) and quality metrics and indicators for analytics of assessment design (Ringtved et al., 2017).

2 PURPOSE OF THE WORKSHOP

The main objective of this workshop is to share and evaluate a range of systems and tools that link learning analytics with learning design in order to provide feedback to developers and identify new insights for future development in the area. The workshop will be very interactive in nature with participants having a chance to use and explore these systems and tools in order to provide feedback and engage in discussions about what we can learn from such developments. Developers of learning analytics/learning design systems and tools will be given the opportunity to submit proposals for their work to be included and reviewed as part of the workshop and can benefit from the feedback provided by fellow participants on their work. Submissions on use case scenarios are also invited as test cases for the submitted systems and tools. The aim is to create a constructive conversation about the ways learning analytics and learning design can be connected through the consideration of current initiatives also helping to identify opportunities and challenges in moving this area forward.

3 WORKSHOP DESIGN

The full-day workshop is expected to attract approximately 30-40 participants and will be open to anyone with an interest in learning analytics and learning design (e.g. teachers, learning designers, researchers, developers, etc.).

3.1 Pre-workshop planning

A website will be established to promote the workshop including information on the purpose and structure of the workshop. This website will also be used to disseminate the outcomes of the workshop as a record for interested members of the learning analytics/learning design community. A call for participation will be released on the 30th October via the website, through various social media channels (e.g. Twitter), and through mailing lists of relevant professional organisations (e.g. SoLAR, Australasian Society for Computers in Tertiary Education (ASCILITE), Association for Learning Technology, etc.). Those individuals/teams who would like their system/tool to be workshopped will be asked to prepare a test system proposal of not more than 1500 words outlining the purpose of the system/tool, the learning design framework that informs it, the design principles of the system/tool, and any aspects on which they would like to receive specific feedback. Individuals/teams who are interested in submitting a use case scenario are invited to submit a scenario description of not more than 800 words on the learning setting that requires the use of a system to support learning design and learning analytics, and a list of performance criteria that can be used to evaluate the system. These proposals will be reviewed by the workshop organisers and 4-5 systems/tools will be selected to be workshopped on the day, together with 3-4 use case scenarios.

3.2 The WORKshop

Requirements for the workshop space include a computer and projector, as well as tables that can be moved to form small groups for the evaluation activities. All participants will be asked to bring along a laptop computer in order to access the systems/tools and Wi-Fi will be required to allow participants to connect to these systems/tools.

The workshop will begin with an overview of the field and current frameworks provided by the workshop organisers. This will set the scene for the conversation that will continue throughout the day around insights for future development and research of the intersection between learning analytics and learning design. It will be made clear to participants that the spirit of this discussion and the evaluation of the systems/tools is to be constructive, not critical.

Following this introduction a review of each of the systems/tools will be undertaken. Presenters will be given 10 minutes to introduce their system/tool and highlight any specific aspects for which they seek feedback. The presenters will be asked to provide an online demonstration environment for the system/tool that workshop participants can then access and explore. Working in small groups, participants will be given an evaluation framework and the use case scenarios that they can use to guide their exploration and consideration of each system/tool. The system/tool development team will be encouraged to circulate the room during this time to help answer any questions groups may have as they explore. At the end of each case scenario a short whole-group discussion will be held to identify key strengths and weaknesses of the system/tools and suggestions for future development. A summary of each groups' feedback will be added to a shared Google Doc that will be made available to all participants at the end of the day.

The workshop will end with a broader discussion facilitated by the workshop organisers of the issues around integrating learning analytics and learning design in light of the experiences participants have had throughout the day. This will allow all participants to learn from the strengths and weaknesses of the systems/tools reviewed, and in doing so, provide insights for useful ways forward.

3.3 Post-workshop dissemination

A summary of the outcomes of the workshop will be made available via the workshop website. Participants in the day will also maintain access to the Google Doc containing the evaluation summaries for each system/tool reviewed. The broader discussion outcomes of the workshop will also form the basis of a discussion paper to be written by the workshop organisers.

REFERENCES

Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gasevic, D., Mulder, R., Williams, D., Dawson, S., Lockyer, L. (2016). A conceptual framework linking learning design with learning analytics. In T. Reiners, B.R. von Konsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (pp. 409-413). New York: ACM.

Bennett, S., Agostinho, S., & Lockyer, L. (2017). The process of designing for learning: understanding university teachers' design work. *Educational Technology Research and Development, 65*(1), 125-145.

Corrin, L., Kennedy, G., de Barba, P., Lockyer, L., Gasevic, D., Williams, D., Dawson, S., Mulder, R., Copeland, S., & Bakharia, A. (2016). *Completing the Loop: Returning Meaningful Learning Analytics Data to Teachers.* Sydney: Office for Learning and Teaching.

Donald, C., Gunn, C., McDonald, J., Blumenstein, M., & Milne, J. (2016). Matching the rhythms of teaching to learning analytics. Paper presented at the *International Consortium of Educational Developers,* University of Cape Town.

Law, N., Li, L., Herrera, L. F., Chan, A., & Pong, T. (forthcoming). A Pattern Language Based Learning Design Studio for an Analytics Informed Inter-Processional Design Community. *Interaction Design and architecture(s).*

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439-1459.

Milligan, S., Ringtved, U. & Corrin, L. (2016). Learning design and feedback processes at scale: Stocktaking emergent theory and practice. *6th International Conference on Learning Analytics and Knowledge*, University of Edinburgh.

Mor, Y., Ferguson, R., & Wasson, B. (2015). Learning design, teacher inquiry into student learning and learning analytics: A call for action. *British Journal of Educational Technology, 46*(2), 221-229.

Persico, D., & Pozzi, F. (2015). Informing learning design with learning analytics to improve teacher inquiry. *British Journal of Educational Technology, 46*(2), 230-248.

Rienties, B., Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: a cross-institutional comparison across 151 modules. Computers in Human Behavior, (60). 333-341.

Ringtved, U., Milligan, S., & Corrin, L., Littlejohn, A., & Law, N. (2017). DesignLAK17: Quality metrics and indicators for analytics of assessment design at scale. *7th International Conference on Learning Analytics and Knowledge*, Simon Fraser University, Vancouver.

Schmitz, M., van Limbeek, E., Greller, W., Sloep, P., & Drachsler, H. (2017). Opportunities and Challenges in Using Learning Analytics in Learning Design. In *European Conference on Technology Enhanced Learning* (pp. 209-223). Springer, Cham.

Learning Design Studio: a Pedagogically Grounded Productivity and Collaboration Platform for Learning Design and Analytics Professionals

Nancy Law
University of Hong Kong
nlaw@hku.hk

Ling Li
University of Hong Kong
lingli.edu@hku.hk

Liliana Farias Herrera
University of Hong Kong
lilians@hku.hk

ABSTRACT: In recent years, there is an increasing recognition of the value of positioning 'teaching as design' (Brown & Edelson, 2003; Goodyear, 2015; Recker et al., 2007) and 'teachers as design professionals' (Laurillard, 2002). Underpinning this recognition is the idea that teaching in the knowledge era should shift from a focus on transmitting knowledge to designing conducive learning environments and experiences to nurture learners' intellectual capacities for 21st century outcomes, grounded on learning sciences-based design principles. In parallel to the developments in learning design are rapid advances in e-learning deployment such as MOOCs, and the application of data analytics and visualization technology to the massive amounts of data generated by learners on online e-learning platforms, particularly on MOOC platforms. In this context, researchers see a great potential in the possible synergy between Learning Design (LD), Learning Analytics (LA) and Teacher Inquiry into Student Learning (TISL), which can together form a virtuous circle for continuous improvements of teaching (McKenney & Mor, 2015; Mor, Ferguson, & Wasson, 2015). It is argued that 'learning analytics offers a powerful set of tools for teacher inquiry, feeding back into improved learning design. However, the promises of LA to improve teaching and learning have largely not been realized for various reasons, including teachers' lack of understanding of LA (Corrin et al., 2016).

Keywords: Learning design, learning analytics

1 INTRODUCTION (PURPOSE)

In recent years, there is an increasing recognition of the value of positioning 'teaching as design' (Brown & Edelson, 2003; Goodyear, 2015; Recker et al., 2007) and 'teachers as design professionals' (Laurillard, 2002). Underpinning this recognition is the idea that teaching in the knowledge era should shift from a focus on transmitting knowledge to designing conducive learning environments and experiences to nurture learners' intellectual capacities for 21st century outcomes, grounded on learning sciences-based design principles. In parallel to the developments in learning design are rapid advances in e-learning deployment such as MOOCs, and the application of data analytics and visualization technology to the massive amounts of data generated by learners on online e-learning platforms, particularly on MOOC platforms. In this context, researchers see a great potential in the possible synergy between Learning Design (LD), Learning Analytics (LA) and Teacher Inquiry into Student Learning (TISL), which can together form a virtuous circle for continuous improvements of teaching (McKenney & Mor, 2015; Mor, Ferguson, & Wasson, 2015). It is argued that 'learning analytics offers a powerful set of tools for teacher inquiry, feeding back into improved learning design. However, the promises of LA to improve teaching and learning have largely not been realized for various reasons, including teachers' lack of understanding of LA (Corrin et al., 2016).

The Learning Design Studio (LDS^{HE}) is an online productivity tool and collaboration platform for learning designers and learning analytics practitioners, developed as an integral part of an on-going project titled "An Open Learning Design, Data Analytics and Visualization Framework for E-Learning" (Law et al., 2017). In this DesignLAK18 proposal, we would like to offer the LDS^{HE} system for testing by workshop participants to evaluate its ability to represent the learning designs of courses, with particular focus on the system's ability to identify the type(s) of Learning Analytics tools and visualization displays that would be appropriate for informing the teachers/learning designers/students for different course learning outcomes and pedagogical designs.

2 LDS^{HE}: UNDERPINNED BY A PRINCIPLED DESIGN PATTERN LANGAUGE CONNECTING LD AND LA

In its design conceptualization, the ultimate purpose of LDS^{HE} is to serve as a platform to connect the LD, LA and TISL communities. To do so requires a common design language that can (1) capture well-constructed pedagogical practices and the underpinning learning design principles, as well as specify the necessary learning analytics appropriate for the intended learning outcomes and chosen pedagogy, and (2) be understood by practitioners and researchers in all of the three targeted communities. To achieve this goal, a major part of the R&D effort in this project is to develop a pattern language, which has been greatly inspired by both the outcome-based educational approach (OBE) (Harden, 2002; King & Evans, 1991), and the Alexandrian pattern language (Alexander, 1964, 1979; Alexander, Ishikawa, & Silverstein, 1977). OBE is an approach that organizes an educational system around what is essential for students to achieve at the end of their learning experience (Harden, 2002; King & Evans, 1991). In this approach, course planning starts with the identification of learning outcomes, followed by a backward design of the learning tasks and assessments to achieve the intended learning outcomes. This approach leads to a constructive alignment between learning outcomes, learning tasks and assessments. OBE advocates a shift in the curriculum design focus from subject matter content as prescribed in textbooks or standardized assessments, to the competencies and performance expected of students after completing a course (King & Evans, 1991). The pattern language developed by Alexander focuses on making the design values and successful design features of spaces and buildings at different levels of granularity explicit (Alexander, 1964, 1979). Alexander's (1977) pattern language comprises of design patterns, which are defined as the core of the solutions to recurrent problems and can be used repeatedly without doing things in the exact same way twice.

The LDS pattern language differs from the Alexandrian pattern language in that it provides a formalism (or language) which can be used to construct design patterns at different levels of granularity in learning design (Law et al., 2017) such the number and characteristics of the patterns that can be constructed is not limited to a fixed number as in the case of the Alexandrian pattern language. This pattern language is being extended such that it can be used to specify LA tools and visualizations for LD patterns represented in the pattern language.

3 A BRIEF INTRODUCTION TO LDS^{HE}

In its design conceptualization, the ultimate purpose of LDS^{HE} is to serve as a platform to connect the LD, LA and TISL communities. To do so requires a common design language that can (1) capture well-constructed pedagogical practices and the underpinning learning design principles, as well as specify the necessary learning analytics appropriate for the intended learning outcomes and chosen pedagogy, and (2) be understood by practitioners and researchers in all of the three targeted communities. To achieve this goal, a major part of the R&D effort in this project is to develop a pattern language, which has been greatly inspired by both the outcome-based educational approach (OBE) (Harden, 2002; King & Evans, 1991), and the Alexandrian pattern language (Alexander, 1964, 1979; Alexander, Ishikawa, & Silverstein, 1977). OBE is an approach that organizes an educational system around what is essential for students to achieve at the end of their learning experience (Harden, 2002; King & Evans, 1991). In this approach, course planning starts with the identification of learning outcomes, followed by a backward design of the learning tasks and assessments to achieve the intended learning outcomes. This approach leads to a constructive alignment between learning outcomes, learning tasks and assessments. OBE advocates a shift in the curriculum design focus from subject matter content as prescribed in textbooks or standardized assessments, to the competencies and performance expected of students after completing a course (King & Evans, 1991). The pattern language developed by Alexander focuses on making the design values and successful design features of spaces and buildings at different levels of granularity explicit (Alexander, 1964, 1979). Alexander's (1977) pattern language comprises of design patterns, which are defined as the core of the solutions to recurrent problems and can be used repeatedly without doing things in the exact same way twice.

The LDS pattern language differs from the Alexandrian pattern language in that it provides a formalism (or language) which can be used to construct design patterns at different levels of granularity in learning design (Law et al., 2017) such the number and characteristics of the patterns that can be constructed is not limited to a fixed number as in the case of the Alexandrian pattern language. This pattern language is being extended such that it can be used to specify LA tools and visualizations for LD patterns represented in the pattern language.

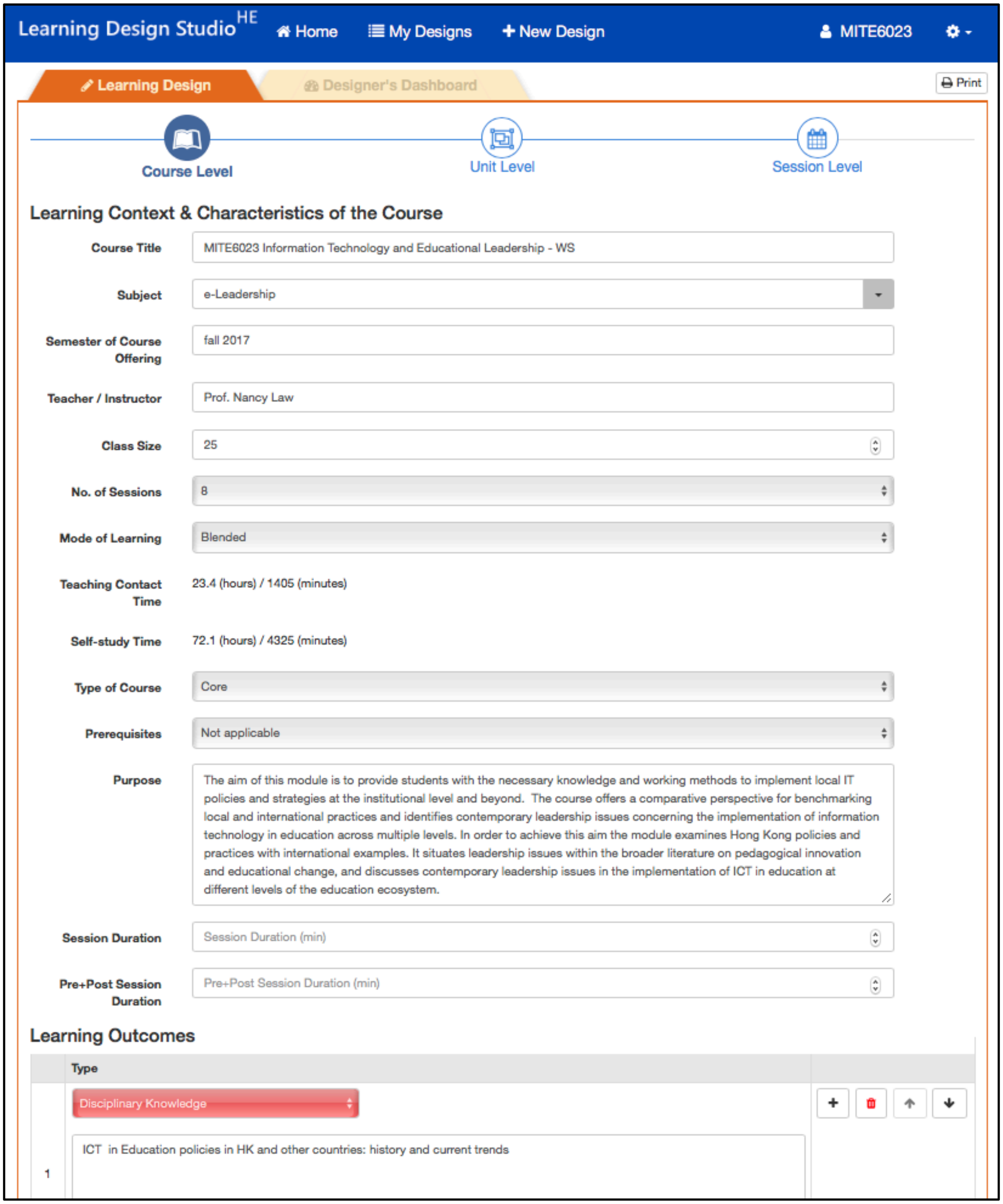


Figure 1. The LDS^HE design interface at the course level.

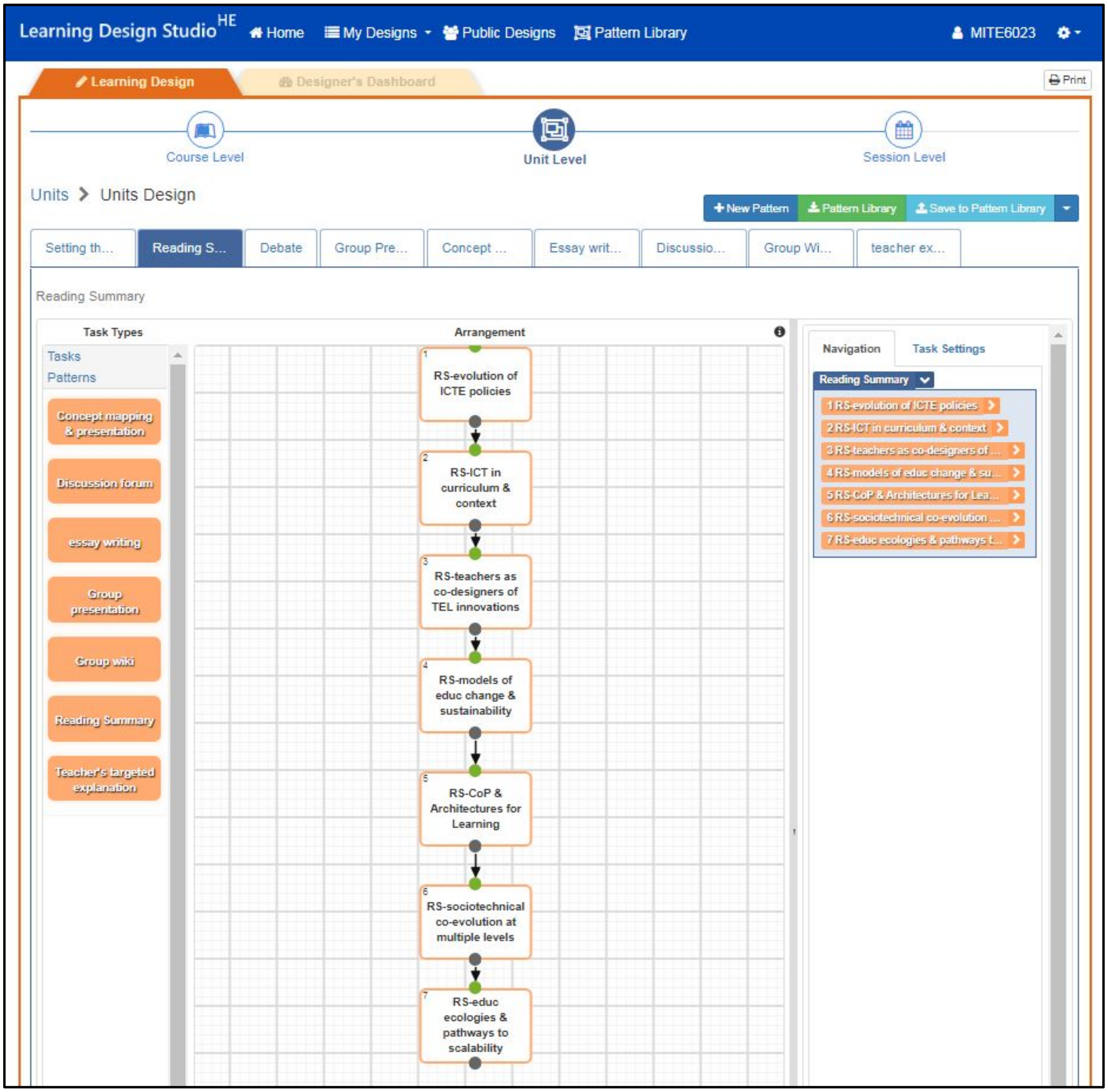


Figure 2. The LDS^{HE} design interface at the unit level.

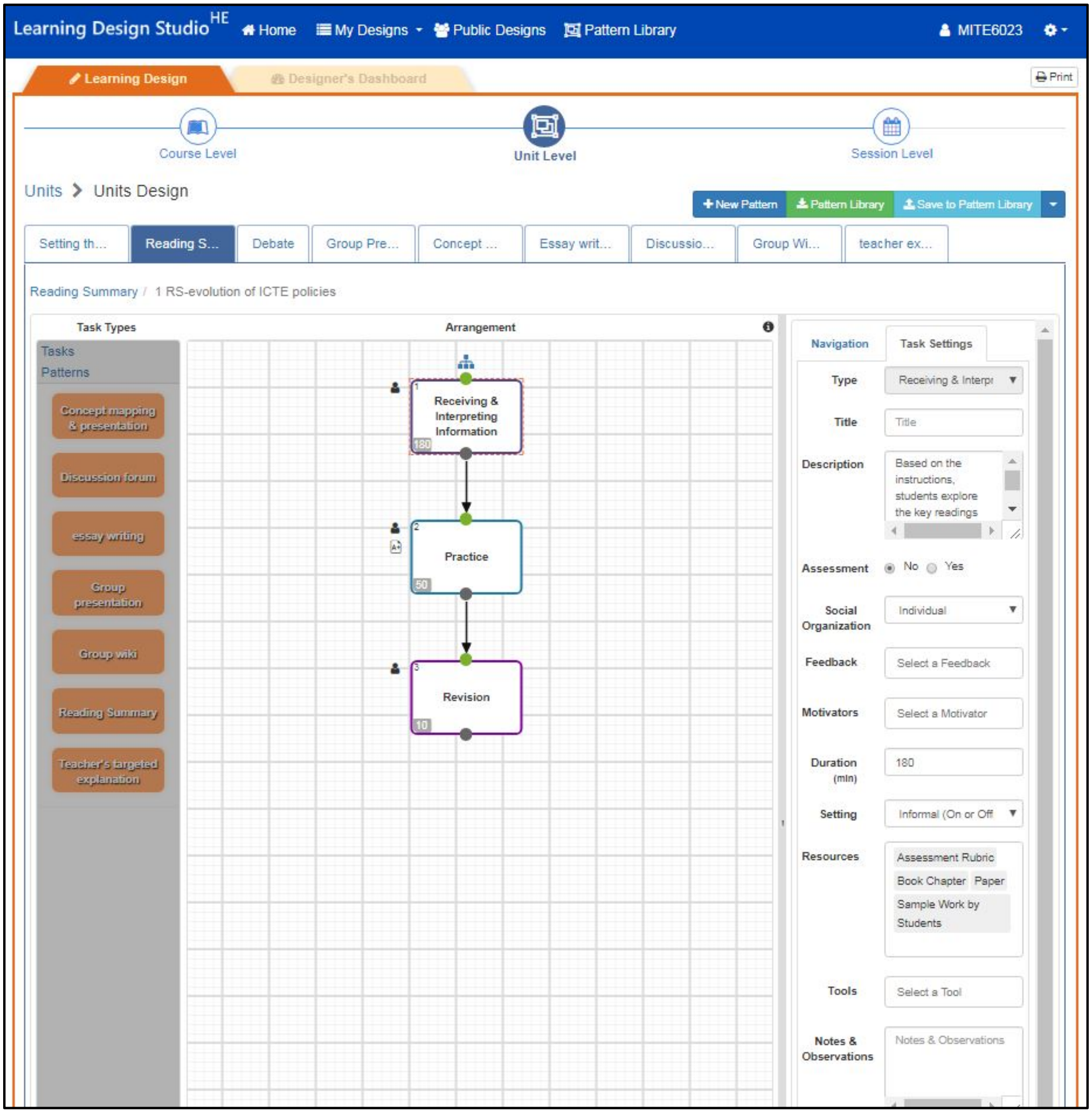


Figure 3. The LDS^HE design interface at the task level.

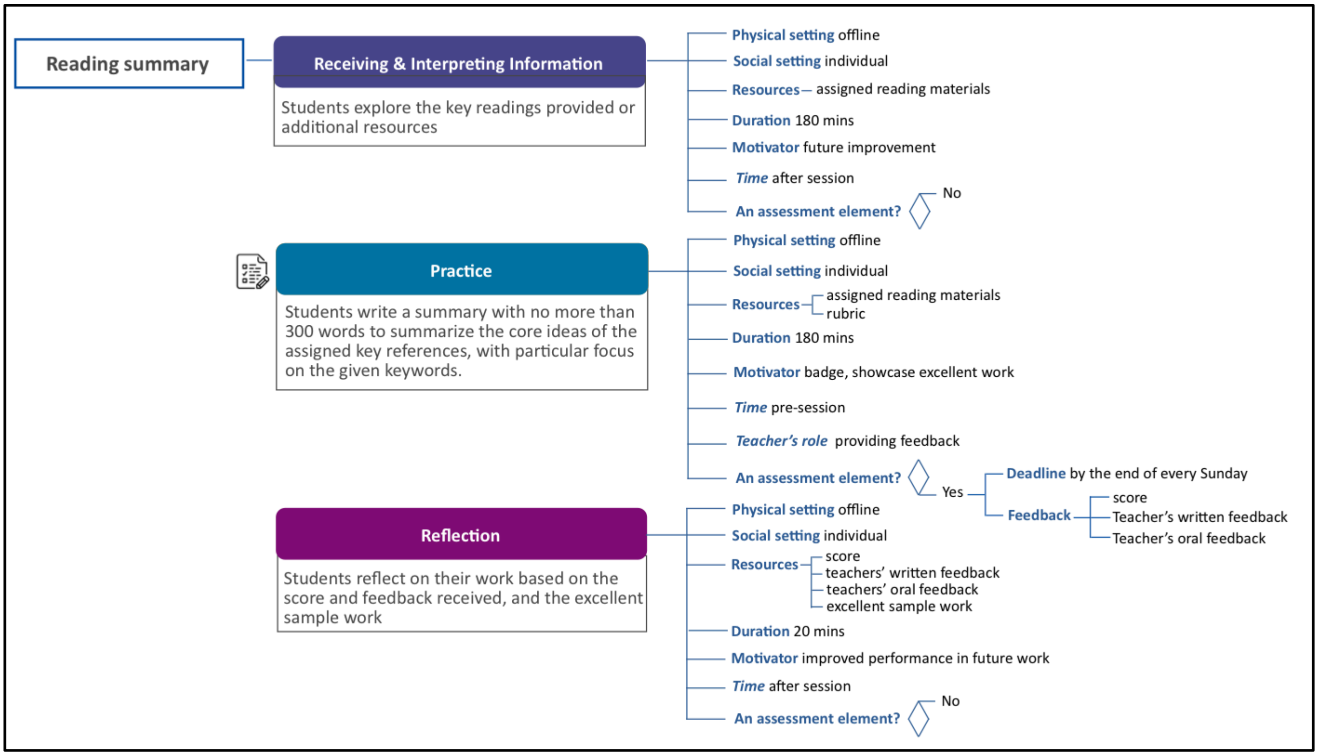


Figure 4. The learning design pattern for the Reading Summary Unit presented in Figure 2.

The hierarchically nested structure of the pattern language and the LDS^{HE} tool provides a systematic and coherent representation of the design elements of a course, with connections within and across different levels of design granularities. This forms a good foundation for connecting different levels of LD with appropriate LA tools and displays, and for the support of teachers' exploration in TISL communities.

4 STRUCTURE OF THE WORKSHOP

In this workshop, we will provide a brief introduction of the theoretical foundations and the features of the LDS. This will be followed by a hands-on session in which participants are invited to try out the LDS^{HE} by representing the selected use case scenarios accepted for this workshop, and to construct specifications of LA functionalities and displays for the provision of different kinds of feedback.

REFERENCES

Alexander, C. (1964). *Notes on the Synthesis of Form*. Harvard University Press.

Alexander, C. (1979). *The Timeless Way of Building*. Oxford University Press.

Alexander, C., Ishikawa, S., & Silverstein, M. (1977). *A Pattern Language: Towns, Buildings, Construction*. OUP USA.

Brown, M., & Edelson, D. (2003). *Teaching as design: Can we better understand the way in which teachers use materials to we can better design materials to support their changes in practice?* Chicago: Northweatern University.

Corrin, L., Kennedy, G., de Barba, P.G., Lockyer, L., GaševicÅL , D., Williams, D., Dawson, S., Mulder, R., Copeland, S., & Bakharia, A. (2016). *Completing the Loop: Returning Meaningful Learning Analytic Data to Teachers.* Sydney: Office for Learning and Teaching.

Goodyear, P. (2015). Teaching as design. *HERDSA Review of Higher Education, 2*, 27–50.

Harden, R. M. (2002). Developments in outcome-based education. *Medical Teacher, 24*(2), 117–120. https://doi.org/10.1080/01421590220120669

King, J., & Evans, K. (1991). Can we achieve outcome-based education? *Educational Leadership, 49*(2), 73–75.

Laurillard, D. (2002). *Rethinking university teaching: a framework for the effective use of educational technology* (2nd ed.). London & New York: Routledge.

Law, N., Ling, L., Herrera, L. F., & Pong, T. C. (2017). A Pattern Language Based Learning Design Studio for an Analytics Informed Inter-Professional Design Community. *Journal of Interaction Design & Architectures, 33*, 92-112.

McKenney, S., & Mor, Y. (2015). Supporting teachers in data-informed educational design. *British Journal of Educational Technology, 46*(2), 265–279. https://doi.org/10.1111/bjet.12262

Mor, Y., Ferguson, R., & Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action. *British Journal of Educational Technology, 46*(2), 221–229.

Recker, M., Walker, A., Giersch, S., Mao, X., Halioris, S., Palmer, B., … Robertshaw, M. B. (2007). A study of teachers' use of online learning resources to design classroom activities*. New Review of Hypermedia and Multimedia, 13*(2), 117–134. https://doi.org/10.1080/13614560701709846.

PerspectivesX: A Tool to Scaffold Collaborative Learning Activities within MOOCs

Marco Lindley, Aneesha Bakharia, Wendy Chalmers, Carrie Finn
UQx, ITaLI, The University of Queensland, Brisbane, Australia
marco.lindley, a.bakharia1, w.chalmers, c.finn@uq.net.au

Linda Corrin
Williams Centre for Learning Advancement,
The University of Melbourne, Australia
linda.corrin@unimelb.edu.au

ABSTRACT: In this workshop, we introduce the PerspectivesX tool which aims to scaffold collaborative learning activities within MOOCS. The PerspectivesX tool is designed to promote learner knowledge construction and curation for a range of multi-perspective elaboration techniques (e.g. SWOT analysis and Six Thinking Hats). The PerspectivesX tool stores learner submissions in a searchable knowledge base which is able to be persisted across course re-runs and promotes the use of natural language processing techniques to allow course moderators to provide scalable feedback. In this workshop we will outline the purpose of this structured collaborative learning tool, and the design principles adhered to during implementation. We also outline key points for discussion at the workshop which relate to embedding analytics to enable instructor feedback and improve linkage to learning design.

Keywords: computer supported collaborative learning, massive open online courses, learning tools interoperability, knowledge construction, critical thinking, idea generation

1 INTRODUCTION

Collaborative learning in MOOCs is predominantly implemented through discussions forums. Research has shown that learners that actively contribute to the course discussion forum are more likely to complete the course and achieve higher grades (Corrin, de Barba, & Bakharia, 2017). A high percentage of learners however, don't engage in a course discussion forum with estimates of forum participation being between 5-10% of all learners (Hill, 2013). The gap between the effectiveness of unstructured collaborative learning tools like forums and their level of student engagement compared to other MOOC instructional content (i.e. videos, quizzes, poll, etc), highlights the need for tools that are able to scaffold collaborative learning activities. PerspectivesX has been developed to bridge this gap.

2 THE PERSPECTIVESX TOOL

The PerspectivesX tool has been designed to scaffold a range of multi-perspective elaboration activities (e.g., SWOT analysis, Six Thinking Hats, etc). It is designed to promote active participation in collaborative learning from all users including learners that either do not participating in discussion forums or that are passive forum participants (i.e., only reading forum posts). The tool encourages learners to contribute and allows them to explore, review and curate submissions from other learners. In a PerspectivesX activity, learners must think about a problem from an assigned or selected perspective and actively contribute their ideas to a knowledge base that is available to all course participants. Instructors can enable an optional curation layer that requires learners to collate ideas from fellow learners in order to complete the remaining perspectives of the activity. Curation is an important feature of the tool. Curation is a 21st century digital literacy that is able to facilitate the development of a learner search and evaluation strategies as well as promote critical thinking, problem solving and participation in networked conversations. (O'Connel, 2012).

The options available to an instructor for setting up a PerspectiveX activity is shown in Figure 1. The tool supports a range of pre-defined idea generation and multi-perspective activities such as Strengths, Weakness, Opportunities and Threats (SWOT) analysis, Six Thinking Hats (de Bono, 2000), Fishbowl (Miller & Benz, 2008) and SCAMPER (Ederle, 1996). PerspectivesX allows instructors to create their own activity template (i.e. define each of the perspectives for learner contribution), shown in Figure 2.

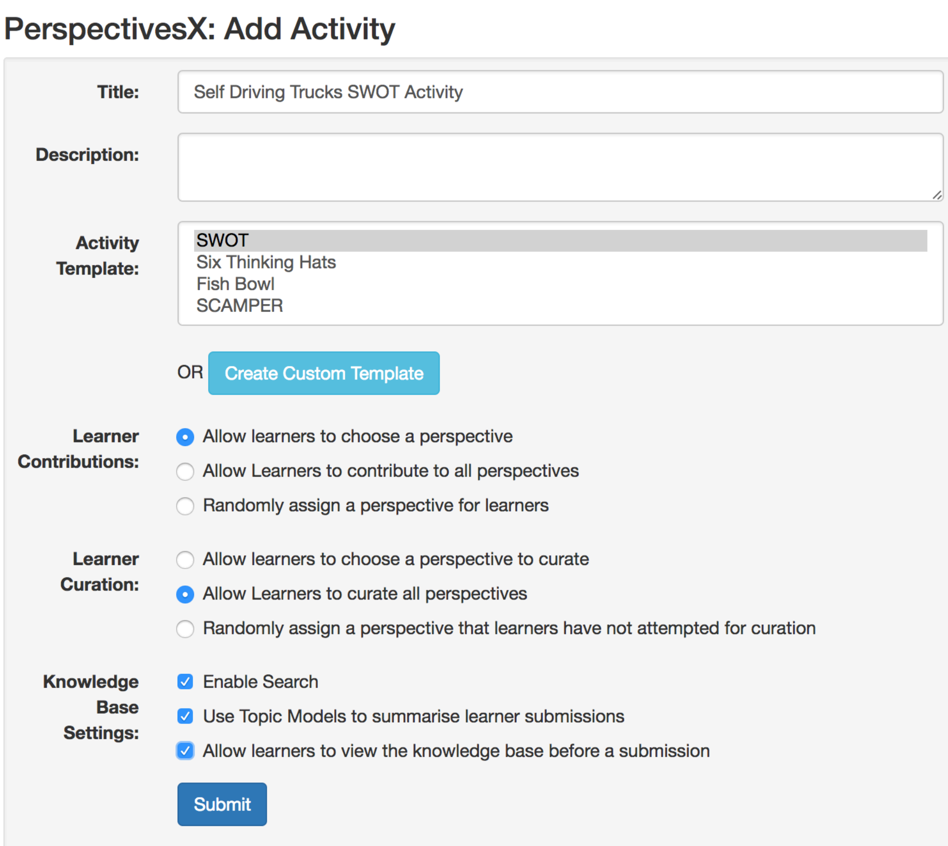


Figure 1: The instructor multi-perspective activity creation interface

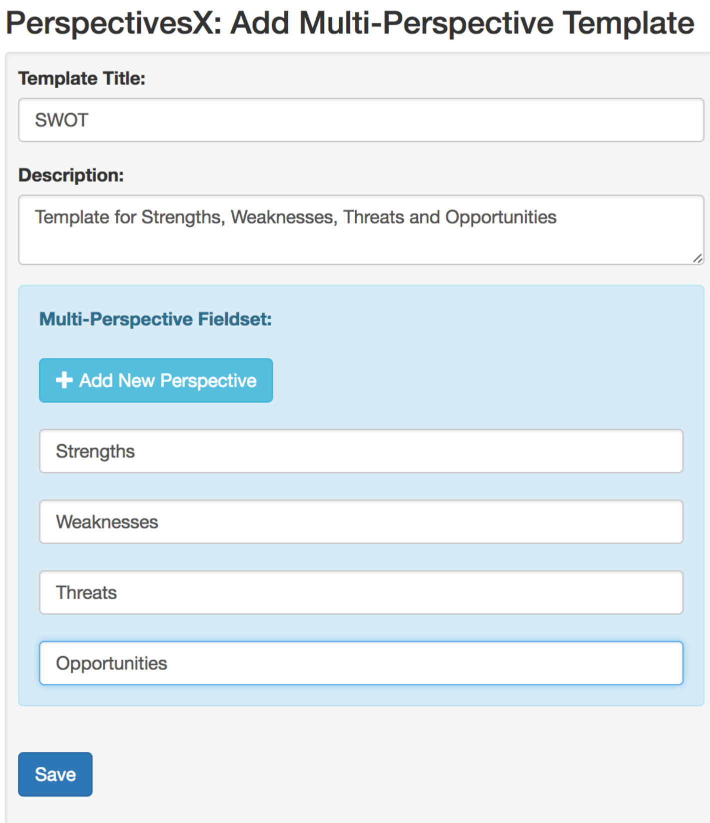


Figure 2 The multi-perspective template creation interface

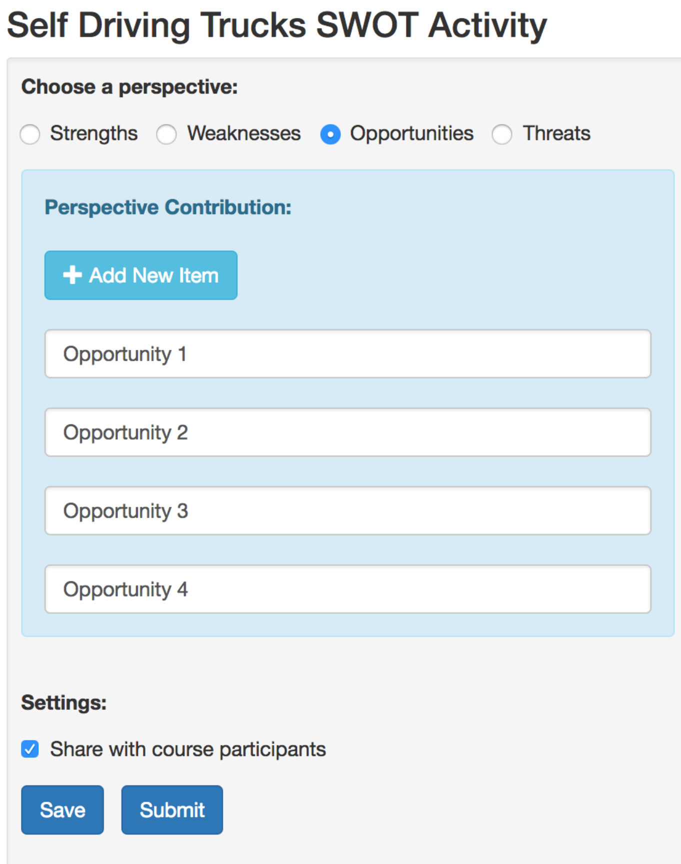


Figure 3 The learner contribution interface (no contribution or curation)

3 LEARNING DESIGN FRAMEWORK

PerspectivesX incorporates ideas from Computer Supported Collaborative Learning (CSCL) and the Knowledge Community and Inquiry model (KCI) (Slotta & Najafi, 2013). KCI uses Web 2.0 tools to add a layer of collective knowledge building to scripted learning activities. PerspectivesX follows the 4 main principles of KCI through its Knowledge base of student answers (principle 1), perspective submission and curation mechanics (principle 2&3) and teacher's moderation of the PerspectivesX content (principle 4) (Slotta & Najafi, 2013).

Tools that support pedagogical scripting of CSCL often employ a visual flowchart metaphor (Haklev, Faucon, Hadzilacos, & Dillenbourg, 2017), providing a high-level overview of the activity. PerspectivesX takes a declarative approach to the configuration of activities encapsulated in a simple form to streamline the configuration of activities. This interface (Figure 1), allows instructors to choose or create an activity template and specify the configuration settings as well as specifying how learners contribute to the perspectives (submission and curation) in an activity.

4 DESIGN PRINCIPLES

The design principles that underpin the PerspectivesX tool are outlined below:

– **Support the design of structured knowledge construction, critical thinking and multi-perspective elaboration activities**

Instructors can design activities that are able to collate structured responses/submissions from learners. The types of responses required by learners should be flexible and allow learners to submit multiple free text responses, media artifacts (e.g., images, infographics, slides, videos, etc) and links to external resources (e.g., website links). Within multi-perspective activities the instructor is able to design activities that allow the learner to select a perspective or be randomly assigned to perspective.

– **Support opt-in and anonymous learner knowledge sharing**

Learners should not be forced to share their submissions with other course participants. Between 5-10% of learners are active discussion forum participants in a MOOC while a larger percentage of learners read forum posts (i.e. passive participation). Many learners may not feel confident making their submissions available to other learners in a non-anonymous environment. Submission should be mandatory in order to receive a participation grade, but learners are able to opt-out of sharing or choose to share anonymously.

– **Support instructor moderation**

Course moderators required the ability to review and curate useful learner contributions. Curated content will help learners to focus their attention on relevant and important submissions from other learners. The learner should be able to view moderator highlighted content in an accessible and intuitive manner. This will allow moderators to use learner submitted work as a starting point to trigger active participation in a discussion forum.

– **Support learner curation**

The scripted collaborative activity should allow for the inclusion of a learner curation sub-activity. As an illustrative example, the collaborative activity might require the learner to submit a single section of a SWOT activity (e.g., strengths) and then at a later stage, curate content from other course participants for the other sections (e.g., weaknesses, opportunities and threats).

– **Support temporal independence**

Both paced and self-paced MOOCs should be able to include scaffolded collaborative learning activities. Learners should be able to contribute to the activity at any time as well as review and curate the submissions of other learners in a time independent manner. This is particularly important for self-paced MOOCs where learners can commence a course at any time and as a result would engage in collaborative learning activities at different times. Discussion forums within self-paced MOOCs are also less active, giving learners limited opportunities to either actively or passively participate.

– **Support knowledge base growth across course re-runs**

Learner contributions should collectively form a knowledge base which becomes available across course re-runs offered in a variety of delivery modes (i.e., paced and self-paced). Initial course runs often have a higher number of enrolled learners and more discussion forum activity as a result. Each MOOC re-run, begins with a refreshed discussion forum which results in community knowledge between courses being lost. Retaining student contributions will facilitate knowledge growth but also poses information retrieval problems. The interface used to display learner contributions will need to therefore include intuitive navigation, free text and tag based (i.e., folksonomy) search functionality.

– **Facilitate the delivery of customised scalable feedback**

While various Natural Language Processing (NLP) and Deep Learning algorithms exist, the ability to accurately grade and provide feedback for free text student submissions within MOOCs has not been realised. There are however techniques that can be used scale feedback provided by instructors, moderators and tutors. These techniques rely on the similarity between learner submissions and can cluster similar learner responses together. Topic modeling using the Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) algorithm is a promising document clustering technique that can be used to find common topics in student submissions. Instructors, moderators and tutors can then view a summary of the topics that exist in learner submissions and provide feedback. Various implementations of using clustering to provide feedback at scale have been discussed by Mohler and Mihalcea (Mohler & Mihalcea, 2009).

5 FEEDBACK& IDEAS.

We welcome feedback on the design principles of PerspectivesX as well as the currently implemented version of PerspectivesX. The source code freely available on Github, https://github.com/UQ-UQx/PerspectivesX. PerspectivesX is implemented as an LTI extension for Learning Management Systems and uses Python, Django and React.

We are particularly interested in discussing how analytics can be embedded within the tool to facilitate customised scalable feedback. We are currently researching a content diversity recommendation system that takes into consideration learner contribution and curation patterns.

REFERENCES

Corrin, L., de Barba, P., & Bakharia, A. (2017). Using learning analytics to explore help-seeking learner profiles in MOOCs. *LAK '17 Proceedings of the Seventh International Learning Analytics & Knowledge Conference.* Vancouver.

de Bono, E. (2000). *Six thinking hats (Rev. and updated ed.).* London: Penguin.

Ederle, B. (1996). *Scamper: Games for imagination Development.* Prufrock Press Inc.

Hill, P. (2013, September 16). *MOOC Discussion Forums: barrier to engagement?* Retrieved from e-Literate: https://mfeldstein.com/mooc-discussion-forums-barriers-engagement/

Miller, R. L., & Benz, J. J. (2008). Techniques for Encouraging Peer Collaboration: Online Threaded Discussion or Fishbowl Interaction. *Journal of INstructional Psychology, 35(1)*, 87-93.

O'Connel, J. (2012). Content curation is the new black. *Collected magazine (6), 4-6*.

Slotta, J., & Najafi, H. (2013). Supporting Collaborative Knowledge Construction with Web 2.0 Technologies. In C. Mouza, & N. Lavigne, *Emerging Technologies for the Classroom* (pp. 93-112). New York: Springer.

Learning Analytics Infrastructure for Seamless Learning

Brendan Flanagan
Academic Center for Computing and Media Studies, Kyoto University, Japan
Flanagan.brendanjohn.4n@kyoto-u.ac.jp

Hiroaki Ogata
Academic Center for Computing and Media Studies, Kyoto University, Japan
hiroaki.ogata@gmail.com

ABSTRACT: Seamless learning offers the opportunity to learn in different environments regardless of location or time. It can also provide insights for teachers into how learning is being conducted in informal situations outside the classroom. Previous work into the analysis of seamless learning has mainly focused on purpose build specialized systems that provide an environment for a specific task. However, as the field of learning analytics matures, we are increasingly seeing the development of modular systems that can be linked together by standards based protocols. This paper proposes the integration of the SCROLL system into a wider modular system to increase the possibilities of seamless learning analytics to inform blended learning design. The proposed system addresses fundamental problems, such as the protection of user privacy and authentication while increasing the availability of data for analysis from other learning systems. Data is collected and stored centrally in a unified form that provides the ability to analyze and visualize learning across numerous environments and contexts.

Keywords: Seamless learning, formal/informal learning analytics

1 INTRODUCTION

Seamless learning offers the opportunity to learn in different environments regardless of location or time. Previous work into the analysis of seamless learning has mainly focused on purpose built specialized systems that provide an environment for a specific task (Mouri, 2017). However, as the field of learning analytics matures, we are increasingly seeing the development of modular systems that can be linked together by standards based protocols. This paper outlines the integration of a seamless learning system called SCROLL into a modular learning analytics platform with the purpose of providing seamless learning analytics spanning numerous learning environments and contexts. We anticipate that seamless learning analytics will be able to provide greater insight into how learning occurs in different contexts and help learners and teachers "connect the dots" between when and in what context students have learnt, revisited, and reflected on knowledge. This paper will focus on the learning of vocabulary as it is the main target of the SCROLL system, however there is potential for the same system and analysis to be applied to a different domain.

2 OVERVIEW OF THE PROPOSED INTEGRATED SYSTEMS

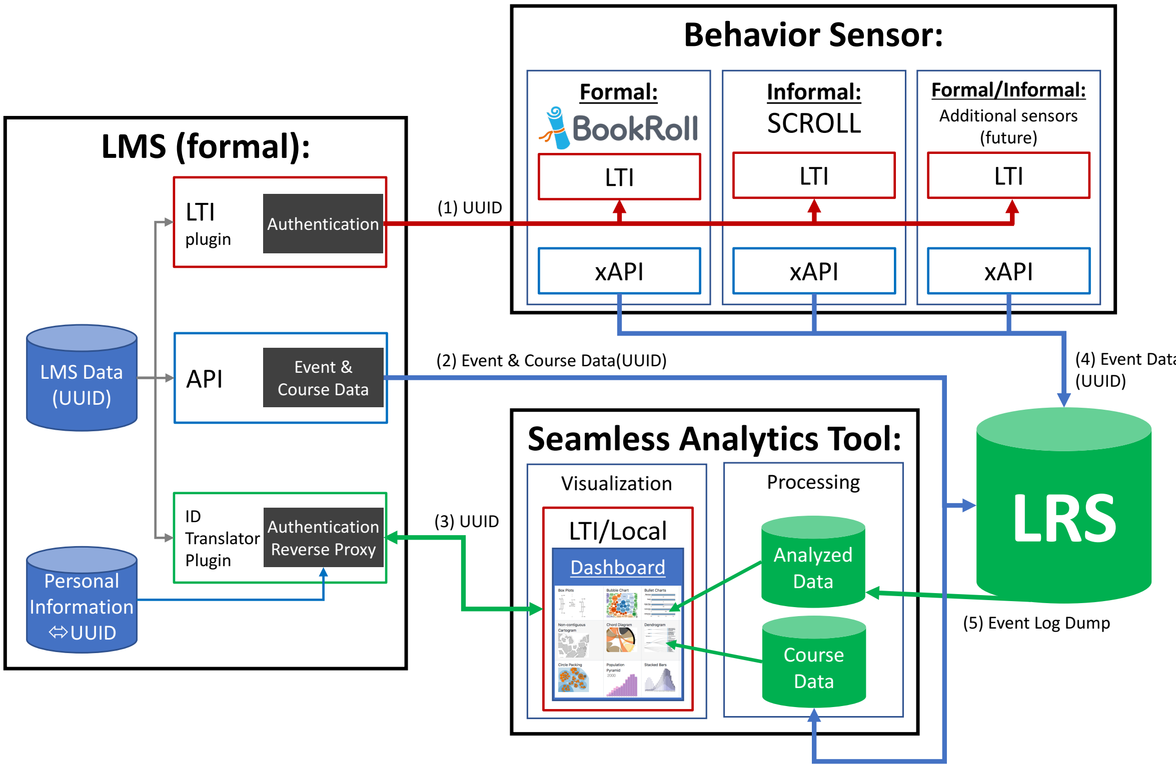


Figure 1: An overview of the proposed learning analytics system.

2.1 Learning Management System

In recent years, several interfaces have been proposed to allow the seamless and secure integration of external tools to augment existing LMS experiences. Some of these interfaces have been proprietary and thus limited the tools that can be integrated. IMS Global Learning Consortium (2016) published the Learning Tools Interoperability (LTI) standard for defining the process of connecting two systems, and how users will transition across these systems without having to authenticate once again with the destination system. During the LTI transition process, information about the user and the context in which the external tool was launched can be transferred from the source system to the target system. Most modern LMS utilize an internal universal unique identifier (UUID) to which personal information, such as: real name and email address are attributed. As shown in Figure 1, we propose that (1) UUID should be transferred to external systems to reduce personal information (a requirement of some education institutions). External tools will then attribute learner events with the LMS's internal UUID that is sent during the LTI launch process, and (4) Event data from behavior sensors and (2) course and event data from the LMS is collected in the LRS (Learning Record Store).

2.2 Behavior Sensors

In the proposed system, user behavior events will be captured by specialist tools that are linked to the LMS by LTI authentication. The behaviors and actions of learners will be sent by an xAPI interface and collected in a central independent LRS.

2.2.1 BookRoll

Digitized learning materials are a core part of modern formal education, making it an increasingly important data collection source in learning analytics. The reading behavior of students has previously been used to visualize class preparation and review patterns by Ogata et al. (2017). The digital learning material reader can be used to not only log the actions of students reading reference materials, such as textbooks, but also to distribute lecture slides, etc. Contents data can also be exported to the LRS for later analysis.

2.2.2 Informal Learning System(s)

In addition to collecting data on user behavior in formal learning situations, we also plan to deploy the SCROLL ubiquitous learning log system that was reported in Ogata et al. (2011) to collect data on user behavior in informal learning environments. SCROLL can be used to support the sharing and reuse of ubiquitous learning logs that are collected in the context of language learning. The addition of behavior sensors that capture event information outside traditional formal classroom contexts enables the support of research into seamless learning analytics of language learners. As the proposed system will collect data from both formal and informal learning environments, this will enable linking of knowledge learnt in either context in addition to information from the LMS, and could be analyzed to predict and extract behaviors of overachieving and underachieving language learners.

Additional integration of specialized language learning tools, such as: testing and exercise systems for the four major skills: listening, speaking, reading, and writing, into the proposed system would provide further opportunities to analyze in detail the behavior of language learners, however at the time of writing this is beyond the scope of this paper and should be addressed in future work.

2.3 Learning Record Store (LRS)

The LRS is an integral part of the proposed system as it will be a central independent point to collect all event data from both the LMS system and behavior sensors. While we have chosen to adopt xAPI as the mode of transporting events data from other systems to the LRS, this is not a strict limitation. We have decided to deploy the latest version of Apereo Foundation's OpenLRS (Apereo Foundation, 2017), which has the ability to support the storing and querying of event data from both xAPI and Global Learning Consortium's Caliper Analytics API (2015). Data from both interfaces are stored in a unified format within the LRS. The collection of data in an LRS also reduces information silos were data is only stored locally in a number of different modular systems, and has the potential to increase the availability of data for analysis. In the proposed system, we plan to take incremental (5) Event log dumps from the LRS database as seen in Figure 1, and sending it to the Learning Analytics Tool for automated processing.

2.4 Connecting the Dots with Seamless Analytics Tool

The Seamless Analytics Tool will have two main functions within the proposed system: the aggregation and processing of data stored in the LRS from disparate systems (LMS, Behavior sensors, etc), and the linking of this data into a visualization in which a learner can see how they have learnt a skill. As the SCROLL system focuses on the learning of vocabulary in informal situations, the example used in this paper will focus on how this could be implemented across formal/informal learning environments.

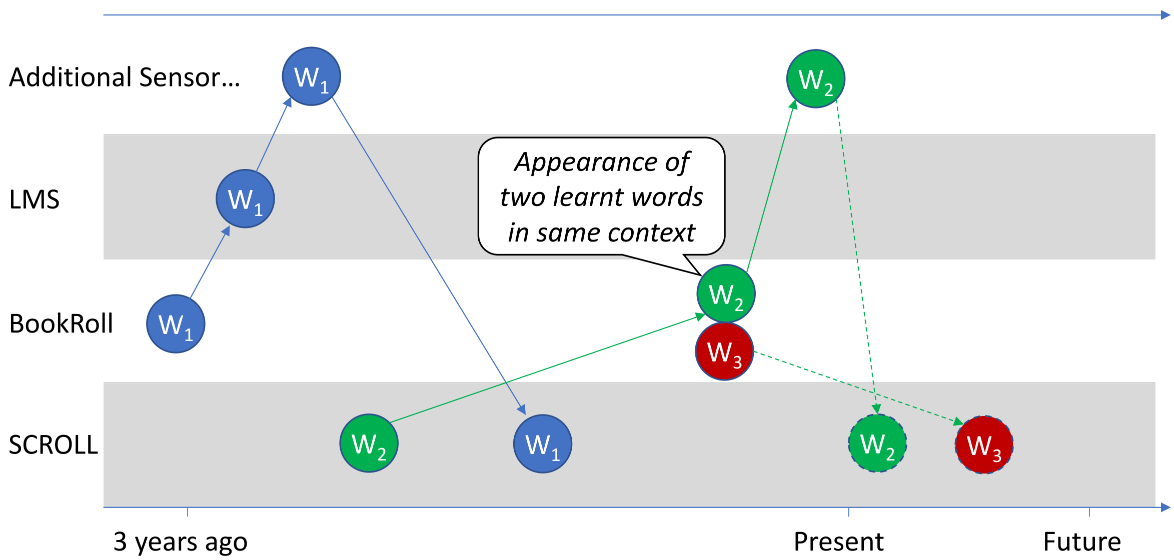


Figure 2: Visualization showing the relation between learnt vocabulary at a personal level.

A mockup of a proposed visualization to inform learners about the relationship between contexts in which they learnt vocabulary is shown in Figure 2. A learner can see that they first learnt the word w_1 while reading a textbook in BookRoll. The same word was also in a quiz they took later on a LMS, and then encounter the word again using an additional sensor system, such as watching a video with transcripts/subtitles in a behavior logging video player. Finally, they encounter and note the word in a real-world context using SCROLL. Aggregated data from other students with similar experiences will be analyzed to predict possible future encounters with known or additional words.

It is proposed that students and teachers will access the portal via a plugin within an LMS that will provide both authentication of the user and also translate the UUIDs that are displayed in the portal into their corresponding real identities depending on their role in the LMS. Teachers who are in charge of a class will be able to view all the student identities of students within that specific class. However, students will only be able to view their own identity, and the identities of their peers will remain anonymous in the results of the analysis. The UUIDs that are displayed in the portal will be marked up with tags to enable quick and effective parsing and translation to real identities by a plugin within the LMS system.

The practice of "connecting the dots" between formal and informal learning is an important part of the language learning process as vocabulary are learnt and reinforced through context (Uosaki et al. 2017). It is anticipated that the use of the propose system will enhance students understanding of knowledge and skills acquired formally by showing related situations that occur in informal learning. A particular example in language learning would be that students can use the system to reflect on the vocabulary learnt in the classroom and how it is being applied in context, location and at what time across disparate learning systems.

3 APPLICATION IN LEARNING DESIGN

The integration of the SCROLL system with a wider variety of learning systems will enable the investigation of learning that occurs across disparate systems that are used in formal and informal contexts. It is anticipated that the proposed system can be used to inform the blended learning design framework. The use of this system could be effective in flipped classrooms, where students can learn formally from learning materials and try to apply the knowledge that they have learnt in an informal context before coming to the class. Experiences of informal learning and application of knowledge can be shared during class time, focusing on reflecting and refining the skills that were acquired before class. This also offers teachers with a unique opportunity to see where and what contexts students are using knowledge and skills. This could then be used to inform the revision of both learning resources and task activities based on use by students outside the classroom.

As the system relies on informal activities being conducted in various location-based contexts, we propose that the system should be assessed on the analysis of prepared pre-class data. In the workshop we would like to evaluate the effectiveness of using the system to inform a flipped class scenario for language learning, and in particular vocabulary in context.

4 CONCLUSIONS

In this paper, we propose the integration of the seamless learning system SCROLL with disparate systems that are currently in use within education institutes, with the purpose of providing a seamless analytics of event data across formal and informal learning contexts. We anticipate this will help learners reflect and "connect the dots" on knowledge they have learnt in various contexts to reinforce their understanding.

REFERENCES

Apereo Foundation. (2017). Apereo Learning Analytics Initiative: Open LRS. (2017). Retrieved September 1, 2017 from http://www.apereo.org/projects/openlrs

IMS Global Learning Consortium. (2015). Caliper analytics. (2015). Retrieved September 1, 2017 from http://www.imsglobal.org/activity/caliper

IMS Global Learning Consortium. (2016). Learning Tools Interoperability (LTI). (2016). Retrieved September 1, 2017 from http://www.imsglobal.org/activity/learning- tools-interoperability

Mouri, K., Ogata, H., Uosaki, N. (2017). Learning analytics in a seamless learning environment. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, 348–357.

Ogata, H., Li, M., Hou, B., Uosaki, N., M El-Bishouty, M., Yano, Y. (2011). SCROLL: Supporting to share and reuse ubiquitous learning log in the context of language learning. Research & Practice in Technology Enhanced Learning 6(2), 69–82.

Ogata, H., Oi, M., Mouri, K., Okubo, F., Shimada, A., Yamada, M., Wang, J., Hirokawa, S. (2017). Learning Analytics for E-Book-Based Educational Big Data in Higher Education. In *Smart Sensors at the IoT Frontier*. Springer, 327–350.

Uosaki, N., Ogata, H., Mouri, K., Choyekh, M. (2017). Implementing Sustainable Mobile Learning Initiatives for Ubiquitous Learning Log System Called SCROLL. In *Mobile Learning in Higher Education in the Asia-Pacific Region*. Springer, 89-114.

Turning the TAP on Writing Analytics

Antonette Shibani
University of Technology Sydney
antonette.shibani@gmail.com

Sophie Abel
University of Technology Sydney
sophie.abel@student.uts.edu.au

Andrew Gibson
University of Technology Sydney/ Queensland University of Technology
andrew.gibson@qut.edu.au

Simon Knight
University of Technology Sydney
sjgknight@gmail.com

ABSTRACT: Writing analytics is seen as a potentially useful technique that uses textual features to provide formative feedback on students' writing. However, for this feedback to be effective, it is important that it is aligned to pedagogic contexts. Such efficient integration of technology in pedagogy could be supported by developing writing analytics literacy. The proposed workshop aims to build this capacity by mapping technical constructs to a broader educational sense for pragmatic applications. It provides a hands-on experience for participants to work with text analytics and discuss its implications for writing feedback. Participants will work with a set of text analysis code to extract features and map them to writing feedback. They will also be given opportunity to develop rules based on extracted text features to write feedback for their own pedagogic contexts.

Keywords: Writing analytics, learning analytics, text mining, writing, writing analytics literacy, data carpentry, hackathon

1 BACKGROUND

Across educational contexts students' written communication is a fundamental concern (National Commission On Writing, 2003; OECD, 2013). Research and commercial tools built on Natural Language Processing (NLP) technologies have gained some traction in supporting the analysis of this writing for educational purposes (e.g., McNamara, Graesser, McCarthy, & Cai, 2014; Shermis & Burstein, 2013). In learning analytics, the sub-domain of *writing analytics* has emerged to support students in their writing practices through the provision of formative feedback that provides insights to educators and students. Two previous workshops (Buckingham Shum et al., 2016; Knight, Allen, Gibson, McNamara, & Buckingham Shum, 2017) have focused on this topic, with community building events focusing on critical perspectives on writing analytics (Buckingham Shum et al., 2016), and the development of greater 'literacy' for writing analytics – building capacity for the use of analytics through its alignment with pedagogic ends (Knight et al., 2017).

This proposed workshop, which would be the third in the series, will build on these previous events by introducing participants to a tool: the Text Analytics Pipeline (TAP). The workshop will build a shared approach to mapping low level textual features to rules that are based on empirical and theoretical work, which can be translated into feedback for students and educators. The LAK17 workshop focused on a need to go beyond simple text analytics, to think about how educators might make effective use of the power of NLP in their teaching. In that workshop the focus was on how a set of existing tools, submitted by participants, map to particular problems in writing, and how each of those tools provide feedback. In the current workshop, participants will work hands-on with text features by generating them using sample code in notebooks, and mapping these features to pedagogic contexts alongside writing useful feedback drawing on the features. At a simple level this might involve, for example, understanding how basic named entity recognition features can be used to develop rules to provide feedback to students: "You've included x, y, and z, but isn't v also an important researcher in this area?". Discussion will focus on technical concerns, and the pedagogic, drawing on research in the pedagogy of teaching writing, and the potential and pitfalls of NLP in addressing that pedagogy. This aligns closely with the theme of user-centered analytics by involving different stakeholders in the design of writing analytics feedback.

For educators to make effective use of writing analytics tools for impact on learning, tools must be integrated into teaching and learning contexts where they guide action, connecting theory, pedagogy, and assessment (Clow, 2012; Knight, Buckingham Shum, & Littleton, 2014; Shibani, Knight, Buckingham Shum, & Ryan, 2017; Wise & Shaffer, 2015; Wise, Vytasek, Hausknecht, & Zhao, 2016). Thus, the third workshop is intended to:

1. Build synergy between writing analytics literacy and writing assessment literacy – that is, build understanding both of the potential of writing analytics and of how to assess writing (both using, and without, analytics tools)
2. Build practitioner capacity for research on their students writing, through developing understanding of how data from writing analytics might provide insights on that writing
3. Build student writing analytics literacy as a means to develop their writing via their critical interaction with text features that contribute to good writing

This workshop will adopt a data carpentry approach to teach writing analytics constructs. Data Carpentry workshops teach beginners the foundation skills needed to conduct data driven research. They are based on Software Carpentry bootcamps (Wilson, 2006), where beginners are taught "basic concepts, skills, and tools for working with data so researchers can get more done in less time and with less pain" (Teal et al., 2015, p. 1). As learners come with varied prior experience, the sessions are an opportunity to build their toolkit so that they can start working with data in their own research. This workshop will guide participants through some vignettes that illustrate the use of a tool (TAP), and its connection to pedagogy, and feedback. The workshop is thus not solely *technical* in nature, instead focusing on developing shared alignment between technical, and social features, towards feedback delivered through analytics tools.

2 SUBMISSIONS AND WORKSHOP FORMAT

2.1 Workshop Objectives

The workshop aims to build writing analytics capacity through developing writing analytics literacy both at this event and beyond. Workshop activities will have two foci: the technical, and the social, both targeted at provision of feedback that supports student writing. The workshop will engage participants with the Text Analytics Pipeline (TAP), and the mapping of textual features output by TAP to rules that can be used to provide feedback to students. The connection between the technical and the pedagogic will be the focus of hands-on activity and will also frame discussion.

2.2 Workshop Activities and Half-day Schedule

This half day workshop will take a participatory approach, blending workshop, tutorial, and hackathon to consider the potential of text analytics tools for supporting writing, and how through use of openly available tools and a data carpentry approach, novices to text analytics – including educators, learning technologists, and others – can be inducted into its potential. The tentative schedule is given below:

Introduction (15 min): Introducing the objectives, presenters and technical set-up for participants.
Introduction to the notebook/workbook (20 mins): The basics of using Jupyter notebooks with some simple text analysis examples.
Tutorial and discussion Part I (50 min): i) Hands-on engagement with text analysis using Jupyter notebooks – extracting features from text.
Break (15 min)
Tutorial and discussion Part II (60 min): i) Hands-on engagement with text analytics – turning features into feedback.
Open discussion (30 min): Discussion and co-creation of a shared resource that maps how other text features can be used in pedagogic applications.
Closing remarks (15 min): Brief summary and discussion of the workshop and future steps.

2.3 Participation, Required Equipment and Dissemination

Participation will be ‘open’ (i.e., any interested delegate may register to attend). The workshop does not require any special equipment (wifi and a room with power strips aside). Participants are encouraged to bring their own devices (laptops best or tablets with keyboards) with a modern web browser. We expect 15-30 participants to attend. An invitation will be extended to participants of previous Writing Analytics workshops to bring different perspectives on the textual features that can be identified and the kinds of feedback that can be provided to help students improve their writing. This workshop will be of interest to a wide range of LAK delegates including: students and researchers actively engaged in writing research, text analytics or writing analytics specifically; educators in schools, universities and businesses; leaders and policymakers; and companies active or potentially active in the field. Some coding skills although not mandatory might be useful.

The workshop organizers are embedded in the learning analytics and related communities. They will make use of listservs (SoLAR, Learning Analytics Google group, EDM-announce, ISLS, SIG-LS, EARLI, ICCE, CHI) and leverage their own personal networks to advertise the workshop. Researchers, practitioners, and funders indicate an increasing interest in writing analytics, and approaches to put writing analytics into practice are currently at the forefront of many learning analytics efforts, thus we anticipate the workshop having popular appeal.

REFERENCES

Buckingham Shum, S., Knight, S., McNamara, D., Allen, L., K. ., Betik, D., & Crossley, S. (2016). Critical Perspectives on Writing Analytics (pp. 481–483). Presented at the 6th ACM Learning Analytics and Knowledge Conference, Edinburgh, UK: ACM. https://doi.org/10.1145/2883851.2883854

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). Retrieved from http://dl.acm.org/citation.cfm?id=2330636

Knight, S., Allen, L. K., Gibson, A., McNamara, D., & Buckingham Shum, S. (2017). Writing analytics Literacy - Bridging from Research to Practice. Presented at the 7th ACM Learning Analytics and Knowledge Conference, Edinburgh, UK: ACM.

Knight, S., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics*, *1*(2). Retrieved from http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3538

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, UK: Cambridge University Press.

National Commission On Writing. (2003). *Report of the National Commission on Writing in America's Schools and Colleges: The Neglected "R," The Need for a Writing Revolution*. College Board. Retrieved from http://www.collegeboard.com/prod_downloads/writingcom/neglectedr.pdf

OECD. (2013). PISA 2015: Draft reading literacy framework. OECD Publishing. Retrieved from http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Reading%20Framework%20.pdf

Shermis, M. D., & Burstein, J. (2013). Handbook of Automated Essay Evaluation: Current Applications and New Directions. Routledge.

Shibani, A., Knight, S., Buckingham Shum, S., & Ryan, P. (2017). Design and Implementation of a Pedagogic Intervention Using Writing Analytics. In W. Chen et al. (Ed.). *Proceedings of the 25th International Conference on Computers in Education*, New Zealand.

Teal, T. K., Cranston, K. A., Lapp, H., White, E., Wilson, G., Ram, K., & Pawlik, A. (2015). Data carpentry: workshops to increase data literacy for researchers. International Journal of Digital Curation, 10(1), 135–143. Retrieved from http://ateson.net/d/49962/index.php/ijdc/article/download/10.1.135/386

Wilson, G. (2006). Software carpentry. Computing in Science & Engineering, 8, 66. Retrieved from http://scitation.aip.org/content/aip/journal/cise/8/6/10.1109/MCSE.2006.122?crawler=true

Wise, A. F., & Shaffer, D. W. (2015). Why Theory Matters More than Ever in the Age of Big Data. Journal of Learning Analytics, 2(2), 5–13. Retrieved from http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4677

Wise, A. F., Vytasek, J. M., Hausknecht, S., & Zhao, Y. (2016). Developing Learning Analytics Design Knowledge in the "Middle Space": The Student Tuning Model and Align Design Framework for Learning Analytics Use. *Online Learning*, *20*(2). Retrieved from http://olj.onlinelearningconsortium.org/index.php/olj/article/view/783

Developing an evidence-based institutional LA policy

Yi-Shan Tsai
The University of Edinburgh
yi-shan.tsai@ed.ac.uk

Maren Scheffel
Open Universiteit Nederland
Maren.Scheffel@ou.nl

Dragan Gašević
The University of Edinburgh
Monash University
dragan.gasevic@ed.ac.uk

ABSTRACT: This workshop aims to support higher education institutions to develop learning analytics (LA) policies that are context-based and evidence-based. The workshop comprises two sessions. The first session will include a number of presentations to introduce the SHEILA policy framework developed by a cross-European project and research findings that have informed this framework. The second session will invite participants to take part in small groups to reflect on the state of LA adoption in their institutional contexts, and use the SHEILA policy framework to draft an institutional policy that considers key action points for LA adoption and addresses identified challenges. The contribution of the workshop is to increase the scalability and sustainability of LA through policy development.

Keywords: policy, learning analytics, higher education, strategy, ethics and privacy

1 INTRODUCTION

For a number of years now, educational institutions have been collecting, storing and analysing the data traces that students and teachers produced and left behind during interactions with virtual learning environments and other digitally traceable systems. The results of these analyses could be fed back to the learners, teachers and institutional management to inform decisions about learning and teaching, thus closing the four-step learning analytics (LA) cycle: generating data, analysing data, feeding data to learners, and activating interventions (Clow, 2012).

While data is described as "the lifeblood for decision-making" by the United Nations (UN Independent Expert Advisory Group, 2014, p.2), and the interest in using data to devise interventions to improve outputs and outcomes in higher education is considered "at an all-time high" (Desouza & Smith, 2016), Ferguson and others (2016) pointed out that the potential of LA, as it has been identified by research, has not been achieved so far due to various barriers. Among the challenges that inhibited the maturity of LA adoption, the lack of practical guidance (Colvin et al., 2015; Ferguson et al., 2014) and insufficient involvement of coordinated leadership (Arroway, Morgan, O'Keefe, & Yanosky, 2016; Siemens, Dawson,

& Lynch, 2013; Tsai & Gašević, 2017a) have been highlighted repetitively in the literature. In light of this, Ferguson and colleagues (2016) made a suggestion for the European policy that "a careful build-up of research and experimentation, with both practice and policies that have a unified European vision" is needed (p.10). Specifically, Tsai and Gasevic (2017a) advocated for the development of institutional LA policies that have considered an individual institution's own cultural, economic, political and technical contexts, so as to ensure the soundness, effectiveness and legitimacy of LA implementations.

In order to leverage strategic planning to scale up the adoption of LA, Ferguson et al. (2014) and Macfadyen et al. (2014) applied and adapted the RAPID Outcome Mapping Approach (ROMA) to learning analytics contexts. The ROMA model was originally designed by the ODI (Overseas Development Institute) to support policy and strategy processes in the field of international development (Young & Mendizabal, 2009). The model begins with defining an overarching policy objective, followed by six steps designed to provide policy makers with context-based information: 1) map political context, 2) identify key stakeholders, 3) identify desired behaviour changes, 4) develop engagement strategy, 5) analyse internal capacity to effect change, and 6) establish monitoring and learning framework. It is designed to be used iteratively rather than linearly.

The ROMA model has also been adopted by a cross-European project – SHEILA (Supporting Higher Education to Integrate Learning Analytics) – to scaffold their analysis of the adoption of LA among 51 HEIs in Europe (Tsai & Gašević, 2017b), based on which they developed the SHEILA policy framework. The SHEILA project team adapted and extended ROMA by incorporating three key elements – action points, potential challenges, and policy prompts, based on the data collected from direct engagement with various stakeholders. Figure 1 explains the concept and structure of the SHEILA policy framework.

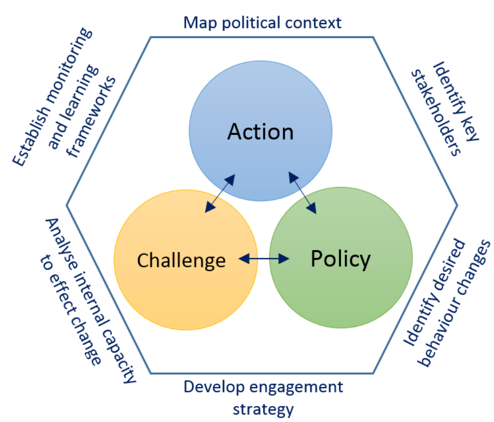


Figure 1: SHEILA policy framework structure

The goal of this workshop is to use the SHEILA policy framework to guide participants to develop a policy draft to increase the scalability and sustainability of LA in their institutions. The workshop is relevant to the meta-issues of the conference – ethics and law, adoption, and scalability. In particular, the workshop will reflect on the needs of different stakeholders and their concerns regarding LA. This is of particular relevance to LAK'18's focus on engaging stakeholders in the design, deployment and assessment of learning analytics. The SHEILA project team held a similar workshop at LAK'17 where they presented findings of their consultations with LA experts and institutional leaders/ decision makers, and guided participants to develop a policy draft using the ROMA model. This year, we will present findings of our consultations with primary stakeholders – teachers and students. We will also showcase the SHEILA policy framework and guide participants to apply it for policy development step by step.

2 WORKSHOP PROGRAMME AND OBJECTIVES

The half-day workshop will be open for anyone interested in institutional policy and strategic planning for LA, particularly those in the following roles in their institutions: policy makers, senior managers and decision makers, LA practitioners and researchers, LA project leaders, data protection and system officers, Information and Technology officers, and academic and student representatives.

The workshop will consist of two main activities – presentations and discussion groups. In the first part of the workshop (1.5h), we will present findings of surveys and focus groups that have been administered to teaching staff and students in four higher education institutions in Europe to understand primary stakeholders' expectations and concerns regarding LA. The first part will also include the presentation of the SHEILA policy framework. The second part of the workshop (1.5h) will engage participants with discussions around the state of adoption of LA in their own institutional contexts, and required policies to ensure effective and responsible implementation. The workshop initiators will use the SHEILA policy framework to guide the discussion process.

The expected number of participants is 30, and the event will be advertised on Twitter, the SHEILA project website (http://sheilaproject.eu/), through the LACE network, and numerous mailing lists. The activities do not require specific equipment besides standard AV.

The goal of this workshop is to assist with the process of developing an institutional policy for the use of LA. There are two main objectives:

1) Participants will discuss and critically reflect on the key action points to take in a systematic adoption of LA, and gain understanding of the potential challenges.

2) Participants will be able to use the SHEILA policy framework to develop a draft of an institutional LA policy that considers LA-related actions and challenges in their institutional contexts.

3 ACKNOWLEDGEMENTS

This work was supported by the Erasmus+ Programme of the European Union [562080-EPP-1-2015-1-BE-EPPKA3-PI-FORWARD]. The European Commission's support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission will not be held responsible for any use which may be made of the information contained therein. The project involved collaborative input from all the partners involved and their contributions are highly appreciated. We would also like to give thanks to our research participants for their valuable contributions.

REFERENCES

Arroway, P., Morgan, G., O'Keefe, M., & Yanosky, R. (2016). *Learning Analytics in Higher Education* (pp. 1–44). ECAR. Retrieved from https://library.educause.edu/resources/2016/2/learning-analytics-in-higher-education

Clow, D. (2012). The Learning Analytics Cycle: Closing the Loop Effectively. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330636

Colvin, C., Rogers, T., Wade, A., Dawson, S., Gasevic, D., Shum, S. B., … Fisher, J. (2015). *Student Retention and Learning Analytics: A Snapshot of Australian Practices and a Framework for Advancement* (pp. 1–91). The Australian Government Office for Learning and Teaching. Retrieved from http://he-analytics.com/

Desouza, K. C., & Smith, K. L. (2016). *Predictive analytics: nudging, shoving, and smacking behaviors in higher education* (EDUCAUSE Rev No. 51) (pp. 10–20). Retrieved from https://er.educause.edu:443/articles/2016/8/predictive-analytics-nudging-shoving-and-smacking-behaviors-in-higher-education

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., … Vuorikari, R. (2016). *Research Evidence on the Use of Learning Analytics: Implications for Education Policy* (JRC Science for Policy Report) (pp. 1–148).

Ferguson, R., Macfadyen, L. P., Clow, D., Tynan, B., Alexander, S., & Dawson, S. (2014). Setting Learning Analytics in Context: Overcoming the Barriers to Large-Scale Adoption. *Journal of Learning Analytics*, *1*(3), 120–144.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D. (2014). Embracing Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Challenge. *Research & Practice in Assessment*, *9*, 17–28.

Siemens, G., Dawson, S., & Lynch, G. (2013). *Improving the Quality and Productivity of the Higher Education Sector: Policy and Strategy for Systems-level Deployment of Learning Analytics*. Sidney, Australia: Society for Learning Analytics Research for the Australian Office for Learning and Teaching. Retrieved from http://www.voced.edu.au/content/ngv%3A64739

Tsai, Y.-S., & Gašević, D. (2017a). Learning analytics in higher education — challenges and policies: A review of eight learning analytics policies. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 233–242). New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3027400

Tsai, Y.-S., & Gašević, D. (2017b). *The State of Learning Analytics in Europe – Executive Summary – SHEILA* (Executive summary). Retrieved from http://sheilaproject.eu/2017/04/18/the-state-of-learning-analytics-in-europe-executive-summary/

Young, J., & Mendizabal, E. (2009). *Helping Researchers Become Policy Entrepreneurs-How to Develop Engagement Strategies for Evidence-based Policy-making* (ODI briefing Papers) (pp. 1–4). London: Overseas Development Institute. Retrieved from http://www.alnap.org/resource/8431

Orchestrating Learning Analytics: Learning Analytics Adoption at the Classroom Level

Luis P. Prieto
Tallinn University, Estonia
lprisan@tlu.ee

María Jesús Rodríguez-Triana
École Polytechnique Fédérale de Lausanne, Switzerland
and
Tallinn University, Estonia
mjrt@tlu.ee

Roberto Martinez-Maldonado
University of Technology Sydney, Australia
roberto.martinez-maldonado@uts.edu.au

Dragan Gasevic
University of Edinburgh, UK
dragan.gasevic@ed.ac.uk

Yannis Dimitriadis
Universidad de Valladolid, Spain
yannis@tel.uva.es

ABSTRACT: The adoption of LA proposals in everyday learning and teaching practice is still slow, and requires effective identification and communication between different stakeholder communities (including researchers, teachers, students and technology developers). To complement high-level institutional, policy-oriented frameworks to promote LA adoption, this workshop proposes to look at how LA innovations impact, or are conditioned by, everyday practice at the classroom level (what some authors call "classroom orchestration"). In this half-day collaborative knowledge building event, participants from these different stakeholder communities bring examples of LA adoption efforts, discuss them through the lens of such classroom orchestration, and further develop frameworks and guidance on what issues should be effectively discussed (and how this communication can be supported).

Keywords: Orchestration, Learning Analytics Adoption, Inter-Stakeholder Communication.

1 BACKGROUND: ORCHESTRATION AND LEARNING ANALYTICS

Despite the recent explosion of research in the field of learning analytics (LA), the adoption of its proposals in everyday classroom practice is still quite limited, and progresses slowly (Ali, Asadi, Gašević, Jovanović, & Hatala, 2013). Multiple researchers, indeed, have looked at the problem of large-scale LA adoption, especially considering how institutions can drive such adoption or develop strategies to favor it (Ferguson et al., 2014; Macfadyen, Dawson, Pardo, & Gaševic, 2014). Many high-level LA adoption frameworks, often aimed at higher education institutions, recognize the need for stakeholder identification and communication with such stakeholders in order to meet their specific needs (Macfadyen et al., 2014). They do not provide, however, concrete guidelines or support for such communication, or what topics should most urgently be addressed by it.

Yet, the problem of slow adoption is not specific to learning analytics, but rather is a manifestation of the more general gap between research and practice that plagues different areas of educational research. Research on systemic and large-scale innovations have noted that success in these endeavors entails a holistic approach that not only considers strategic policy, but also the impact on classroom-level practice (Looi, So, Toh, & Chen, 2011). Especially crucial in this regard is the role of teachers/practitioners as major gatekeepers for the technological and practice innovations that reach the classrooms. The need for proposals that take into account the often-dire contextual constraints of classroom practice, has led to the notion of designing for classroom orchestration (defined as "the process of productively coordinating supportive interventions across multiple learning activities occurring at multiple social levels" (Dillenbourg, Järvela & Fischer, 2009, p. 12).

Recent reviews of technology-enhanced learning (TEL) and LA literature have mapped LA-specific issues and frameworks with this general concern about the impact on classroom-level practice, focusing on the knowledge gaps that arise between the different stakeholders (e.g., teachers, students, researchers and technology providers) during the adoption of an LA tool (Figure 1). In this workshop, such mapping serves as a starting point (or a 'boundary object', Star & Griesemer, 1989) for the dialogue among stakeholders about what issues need to be shared and understood, and how to support more effective inter-stakeholder communication in this process of adoption.

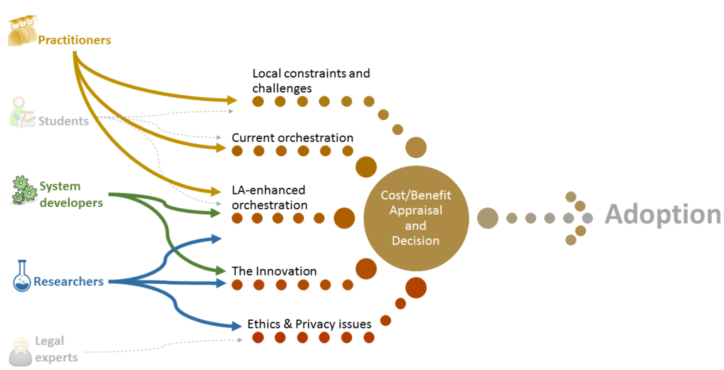


Figure 1: A framework for LA adoption at the classroom level

This workshop builds upon *previous events* that brought together different stakeholders groups to discuss classroom-level factors and how they condition adoption, both in the case of LA[1], and in other fields of educational technology research[2]. As such, this workshop serves as a complement to policy-oriented LA adoption workshops held in previous LAK conferences (Tsai, Gaševic, Muñoz-Merino, & Dawson, 2017).

2 GOALS AND OUTCOMES

The present workshop aims at engaging the LA research community in a dialogue with other stakeholders (practitioners, technology developers), to share concrete adoption experiences and discuss about what factors influence a successful adoption of LA solutions, by looking at the classroom level, rather than a more institutional viewpoint.

The *outcomes* of the workshop may thus consist of an enriched framework and instruments to support communication about the adoption of LA (from the perspective of orchestration). These products, together with the rest of workshop-generated materials will be shared in the workshop website and social media, and eventually provide the basis for a SoLAR sub-community around this topic. Other future steps towards the establishment of this community will also be discussed during the workshop, including a journal special issue, a follow-up workshop in the frame of LAK'19 or other related conferences, or the creation of a virtual community to share experiences and refined/contextualized boundary objects around inter-stakeholder communication for LA adoption. As such, the workshop is strongly aligned with the LAK'18 theme, with respect to LA innovation and adoption in authentic "classroom-level" contexts, promoting stakeholder engagement and communication, accountability, and co-design of effective LA tools.

3 THE WORKSHOP CONTRIBUTIONS

Six contributions make up this workshop proceedings volume. All of them present accounts of LA adoption processes of particular LA innovations or tools, or tease out different factors that will play out in the adoption of such tools in particular settings. The contributions, however, represent the points of view of different stakeholders often involved in LA adoption processes:

1. In their paper Azcona, Hsiao and Smeaton describe the PredictCS system for automated, personalized feedback to university learners of programming skills, from a researcher team perspective. They also describe the stakeholders involved and the ongoing adoption process of this LA innovation in Dublin City University.

2. From a practitioner perspective, Dawkins describes a series of "hacks" he performed to overcome the limitations of the institutional learning management system (LMS) in terms of

---

[1] In the LASI Spain 2016 event: http://lprisan.wixsite.com/orla2016

[2] In CSCL2015 (https://sites.google.com/site/occw15/), or ICLS2012 (https://www.isls.org/icls/2012/program/#workshops)

learning analytics. He also outlines issues unearthed by the adoption process of these LA hacks, which may also be useful for other practitioners and LA researchers/developers.

3. In their contribution, Liang, Nishimura, Nishimura and Chapa-Martell describe another particular LA tool (LessonSpectrum), which focuses on data analysis and visualization in second language education. The authors also highlight the involvement of multiple stakeholders in the initial phases of adoption of this innovation.

4. De-la-Fuente-Valentín, Zhang, Arakistain, Zheng and Burgos present a case study into a new approach to enable spontaneous teacher inquiries over LA datasets using natural language queries. This approach has been devised from lessons learned after several attempts to have LA tools adopted in practice.

5. De-la-Fuente-Valentín and Burgos, on the other hand, describe a different LA innovation (the A4Learning tool), aimed at supporting tutoring processes in online learning. The description of one of the phases in the adoption (its usage in two authentic courses) helps illustrate several factors that can challenge or support the adoption of this innovation in the future.

6. In the last contribution, Shehata provides a complementary perspective on the evaluation and adoption of a particular LA feature within a larger LMS, told in this case from the (industry) technology developer perspective. The lessons learned from this adoption process highlight issues of inter-stakeholder communication and capacity building.

4 CONCLUSION

During the workshop, we will present and discuss the "Orchestrating Learning Analytics" framework for multi-stakeholder communication on the adoption of LA innovations. The aforementioned contributions will be presented and analyzed using this framework, as examples of LA adoption process. In small-group collaborative activities, participants will further synthesize lessons learned, critical issues in the adoption process, and mechanisms to support inter-stakeholder communication during such (often iterative) adoption.

We encourage the reader to head to the workshop website (https://sites.google.com/view/orla-ws-2018/home) to learn about the outputs of this multi-stakeholder community event.

ACKNOWLEDGEMENTS

This work has been partially funded by the European Union in the context of CEITER and Next-Lab (Horizon 2020 Research and Innovation Programme, grant agreements no. 669074 and 731685), the Spanish Ministry of Economy and Competitiveness (projects TIN2014-53199-C3-2-R and TIN2017-85179-C3-2-R), the Regional Government of Castilla y León (project VA082U16), and the Spanish Ministry of Education and Culture (PRX17/00410).

REFERENCES

Ali, L., Asadi, M., Gašević, D., Jovanović, J., & Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education*, *62*, 130–148. doi:10.1016/j.compedu.2012.10.023

Dillenbourg, P., Järvelä S. & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In Technology enhanced learning. Springer, 3–19.

Ferguson, R., Macfadyen, L. P., Clow, D., Tynan, B., Alexander, S., & Dawson, S. (2014). Setting Learning Analytics in Context: Overcoming the Barriers to Large-Scale Adoption. *Journal of Learning Analytics, 1*(3), 120–144. doi:10.18608/jla.2014.13.7

Looi, C. K., So, H. J., Toh, Y., & Chen, W. (2011). The Singapore experience: Synergy of national policy, classroom practice and design research. *International Journal of Computer-Supported Collaborative Learning, 6*(1), 9-37.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D. (2014). Embracing Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Challenge. *Research & Practice in Assessment, 9*, 17-28.

Star, S L, & Griesemer, J. R. (1989). Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3), 387–420. doi:10.1177/030631289019003001

Tsai, Y. S., Gaševic, D., Muñoz-Merino, P. J., & Dawson, S. (2017). LA policy: developing an institutional policy for learning analytics using the RAPID outcome mapping approach. In *Proceedings of the Seventh LAK Conference* (pp. 494-495). ACM.

PredictCS: Personalizing Programming Learning by Leveraging Learning Analytics

David Azcona
Insight Centre for Data Analytics, Dublin City University
David.Azcona@insight-centre.org

I-Han Hsiao
School of Computing, Informatics & Decision Systems Engineering, Arizona State University
Sharon.Hsiao@asu.edu

Alan Smeaton
Insight Centre for Data Analytics, Dublin City University
Alan.Smeaton@insight-centre.org

ABSTRACT: This paper presents a new framework to harness sources of programming learning analytics at a Higher Education Institution and how it has been progressively adopted at the classroom level to improve personalized learning. This new platform, called PredictCS, automatically detects lower-performing or "at-risk" students in computer programming modules and automatically and adaptively sends them feedback. PredictCS embeds multiple predictive models by leveraging multi-modal learning analytics of student data, including student characteristics, prior academic history, logged interactions between students and online resources, and students' progress in programming laboratory work, and their progression from introductory to advanced CS courses. Predictions are generated every week during the semester's classes. In addition, students are flexible to opt-in to receive pseudo real-time personalized feedback, which permits them to be aware of their predicted course performance. The adaptive feedback ranges from programming suggestions from top-performers in the class to resources that are suitable to bridge their programing knowledge gaps.

Keywords: Learning Analytics, Machine Learning, Computer Science Education

1 INTRODUCTION

PredictCS is a Predictive Analytics platform for Computer Science courses that notifies students based on their performance using past student data and recommends most suitable resources for students to consult. These notifications have been widely adopted since Purdue University launched Course Signals (Arnold & Pistilli, 2012). Dublin City University's PredictED project emulated Purdue's and notified students their rank within the class (Corrigan & Smeaton, 2015). Both systems yielded impressive improvement in first-year retention rates.

Recently, researchers have been working on augmenting the Integrated Development Environment (IDE) or programming environment by crowdsourcing code solutions. Students are suggested error corrections or solutions that peers have applied before. Java has been the programming language targeted the most with platforms such as University of Durham's BlueFix (Watson & Godwin, 2012). In addition, also using code snapshots, Carnigie Mellon's ITAP (Intelligent Teaching Assistant for Programming) provides hints as feedback while programming and has been implemented in CloudCoder (Rivers & Koedinger, 2017).

In Dublin City University, Dr. Stephen Blott, lecturer at the School of Computing, has developed a Virtual Learning Environment (VLE) for the teaching of computer programming. The system includes a grading platform that provides real-time feedback on each programming submission by running a suite of test cases. This system gives no code suggestions or personalized help for errors.

2 INFRASTRUCTURE

2.1 Development of the platform

PredictCS has been designed to enhance and personalize the student's learning in programming modules. The platform sits on top of Dr. Blott's VLE grading system and leverages that submission data. Figure 1 shows how PredictCS is inputted a combination of student characteristics, past academic results, programming submissions and interactions with the material to generate predictions and recommendations to the students, lecturers and Faculty.

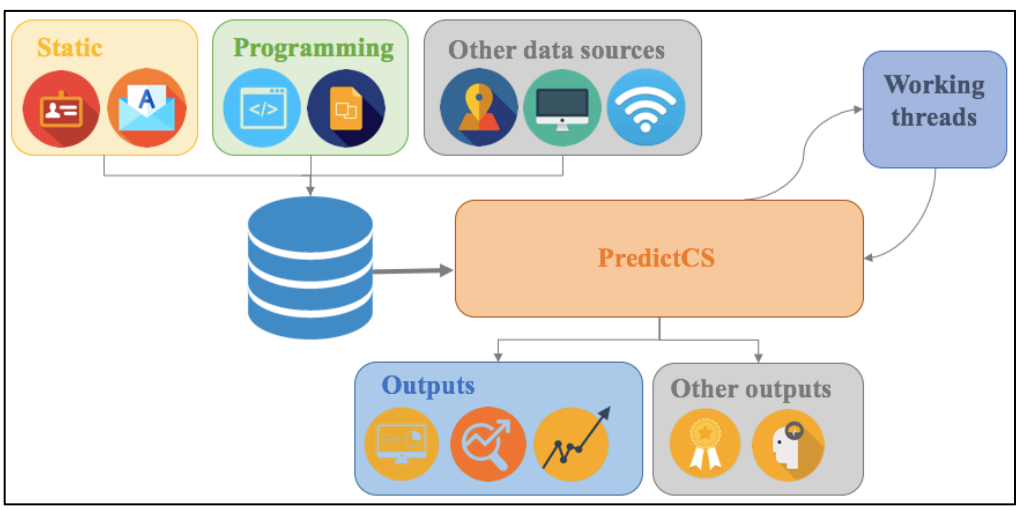


Figure 1: Infrastructure of PredictCS

Our contribution is to combine the detection of students "at-risk" in programming modules and the feedback we provide students with timely weekly notifications by leveraging Learning Analytics approaches jointly with suitable code solutions and resources. Figure 2 shows how the combination of Learning Analytics with programming data enables us to provide students with personalized feedback in Computer Science courses. The feedback contains a message based on their predicted performance, recommended resources and programming code solutions from top-students in the class. Code solutions are found to have a positive effect on the student's learning (Nguyen et al, 2016). We work with submissions and typically short programs so the feedback given to students can be on submission at the platform or via email and we do not have snapshots of students developing the code to learn from their design. We are focusing on email notifications but it can change in the future.

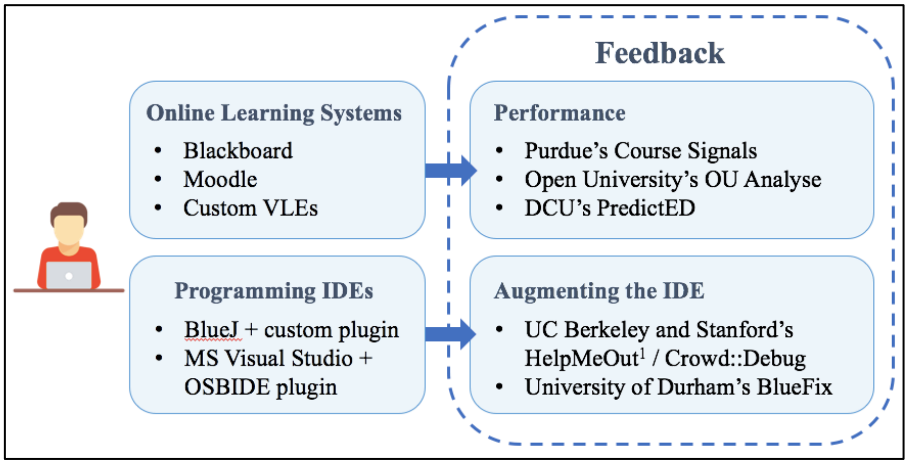


Figure 2: PredictCS combines Learning Analytics with programming data

2.2 Innovation

The platform's goal is to enhance and motivate students to learn programming skills. We apply Learning Analytics as we gather significant data about students learning to program and interacting with the material. Using that digital footprint allows us to understand better their knowledge gaps and help them learn more. We do so by sending notifications about performance, suggested programs and material on a weekly basis to students that opt-in. We developed a recommendation engine as a module for PredictCS that suggests code solutions from top-students in the class as seen in Figure 3. In addition, specific material lessons are recommended to students on these notifications.

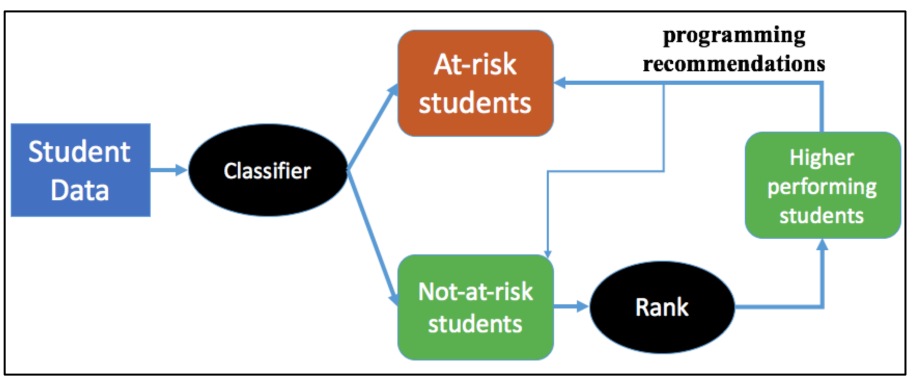


Figure 3: PredictCS Recommendation Engine

The main benefits of the innovation are to help navigate students to the areas they should focus on for their learning progression and, at the same time, notify lecturers where the class as a whole is struggling in order for them to adapt their teaching and curriculum. This can be achieved by using data gathered from past student cohorts and applying machine learning and collaborative filtering techniques.

The ideal scenario for teaching and learning practices would be to design and plan material based on the student's progress and learning using our platform, where lecturers can detect and attend to disengaged students, see how the learning of concepts get propagated through the class and seeing how feedback helps their learning and motivation. We found face-to-face discussions with researchers on the project help lecturers make the most of the platform.

3 STAKEHOLDERS

3.1 Teachers

As a prerequisite for the use of PredictCS, teachers have to utilize the custom grading VLE for the teaching of programming. That requires the teachers to upload the laboratory exercises and the corresponding unit test cases to a server for the grading of them. That way, we gather data about the programming submissions from students. Lecturers should be comfortable uploading and checking material using web technologies, have a basic understanding of probabilities and how PredictCS perform predictions and recommendations. Believing data-driven technologies can enhance and help their teaching is desirable.

PredictCS supports the preparation of teaching activities by Informing lecturers where students are having more issues (i.e. which programming exercises students fail the most) and how students are likely to perform on the next assessment. In short, a report is also sent to lecturers with top issues of the week (the exercises students have the most difficulty) and ranked list of students and their performance likelihood of failing the next assessment. The system sends weekly notifications to lecturers containing who may be at-risk and what are the difficulties students find, but unfortunately, we are not able to measure whether teachers read these emails thoroughly.

PredictCS gives them insights now about how students are doing in the laboratory sessions. Lecturers try to adapt based on the students' needs and issues with the material. They are now better aware of their students' progress by looking at PredictCS and, also, their results at laboratory examinations. The biggest challenge for the lecturers is to find ways to reach as many students as possible in a personalized way. Automatic grading and data analytics tools are helping us to move forward in that space.

Teachers were interviewed at the end of each semester and admitted with large class sizes for programming modules it is practically impossible for them to monitor each student personally and these automated approaches were useful. Simply seeing the list of students marked red or green each week gave them a sense for how things were going. However, they pointed out the predictions seemed too negative as researchers were trying to maximize the students marked "at-risk" while keeping a good balance between the two classes.

3.2 Students

In order to receive personalized feedback, students need to opt-in to the notifications. After their first laboratory exam, a pop-up will appear on the VLE's submission platform to opt-in or out and where they can find further information about the project. Then, the system sends customized notifications every week based on the progress, suggested resources to focus their learning and programming code solutions to the students that opted-in. Every notification contains an unsubscribe link for them to opt-out at any time. Students do not need any pre-requisites but a positive attitude towards problem solving for them to stay motivated.

Researchers gathered the students' opinions about the platform and the notifications via a written questionnaire. Most students would recommend the system to students attending the same module next year or would like to see it included in other modules. Students who were doing very well were getting a similar response every week and were demanding more personalized feedback.

3.3 Researchers

Learning to program involves a variety of complex cognitive activities, from conceptual knowledge construction to basic structural operations. This discipline is very logical and some students find it easier than others in the beginning. Researchers on this project strive to provide the tools for lecturers to adapt their teaching based on their classes' progression and for students to stay engaged and focus where they need to. A web application is also maintained by these researchers where the analysis of the machine learning classifiers, the predictions for every week and all the notifications sent to students can be found. Lecturers and Faculty have access to this platform.

3.4 Faculty

Faculty at Dublin City University's School of Computing have been running research projects around student retention and engagement through Learning Analytics for the past few years. It took them months to prepare a proposal for the University Research Ethics Committee. Now, researchers can request access to this programming related data and static information about students. The University Research Ethics Committee and the data commissioner will review these applications and grant access if appropriate. The school administration owns the data generated.

Dublin City University is partnering with other Higher Education Institutions such as Arizona State University to share methodologies and pioneer approaches around Learning Analytics and Computer Science Education.

4 ADOPTION

In Computer Science Education, lecturers teach programming courses that contain a considerable amount of laboratory work for students to learn these skills. That allows data-driven technologies such as our platform to perform well.

PredictCS is an ongoing project and has been in place for more than three semesters in our university and ran in more than six computer programming modules. At first, we ran a retrospective analysis and developed predictions for student performance using cross-validation techniques on past student data that confirmed the potential of these systems. The following semester we ran pseudo-real time weekly predictions using only one year of training data and targeted a new cohort of students individually during laboratory sessions (Azcona & Smeaton, 2017). Then, the next year, we sent notifications to students based on their performance and suggested resources and have been increasingly adding code solutions, suggested reading material and laboratory sessions to focus their attention on.

Students that corrected their programs and re-submitted them based on the suggestions improved their performance with respect to the ones that did not. In addition, the differential improvement between the two laboratory exams also increased the year the predictions were run and the notifications were sent compared to the previous academic year that our models are trained with.

In terms of the ethical issues, some students may want the data generated by their learning to be deleted and they can request the university to do so. However, that's the way we measure their progress and performance.

5 CONCLUSION

This study has proven to successfully create a programming digital footprint we can leverage. Our data approaches to select a learning algorithm or a subset of predictors on data models enabled us to identify students having issues in programming courses and to provide them with timely interventions. This system and these notifications have shown promising results and will be implemented in more programming modules so we can have a broader impact. We will also be more vocal about the project so more students can opt-in, understand and benefit from this research. We are eager to provide our students with more detailed programming recommendations, suitable material and other actions to fill the knowledge programming holes they may have while learning Computer Science programming design at our university so we are working towards improving some modules of the platform.

REFERENCES

Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267-270). ACM.

Azcona, D., & Smeaton, A. F. (2017, September). Targeting At-risk Students Using Engagement and Effort Predictors in an Introductory Computer Programming Course. In *European Conference on Technology Enhanced Learning* (pp. 361-366). Springer, Cham.

Corrigan, O., Smeaton, A. F., Glynn, M., & Smyth, S. (2015). Using educational analytics to improve test performance. In *Design for Teaching and Learning in a Networked World* (pp. 42-55). Springer, Cham.

Nguyen, Q., Tempelaar, D. T., Rienties, B., & Giesbers, B. (2016). What learning analytics-based prediction models tell us about feedback preferences of students. *Quarterly Review of Distance Education, 17*(3), 13.

Rivers, K., & Koedinger, K. R. (2017). Data-driven hint generation in vast solution spaces: a self-improving python programming tutor. *International Journal of Artificial Intelligence in Education, 27*(1), 37-64.

Watson, C., Li, F. W., & Godwin, J. L. (2012, September). BlueFix: using crowd-sourced feedback to support programming students in error diagnosis and repair. In *International Conference on Web-Based Learning* (pp. 228-239). Springer, Berlin, Heidelberg.

Learning Analytics Classroom Hacks: Examples from an Australian University

Author: Roger Dawkins
Western Sydney University
r.dawkins@westernsydney.edu.au

This study discusses learning analytics solutions I implemented in three blended undergraduate units over two years. My objective was to gather data on students' engagement with content. The solutions are "hacks" because they resolved the limitations of my university's learning management system. I note how these hacks enabled me to make data-driven iterative changes to unit design/content; in addition, student surveys show an increase in satisfaction with the resources of one of the units. I also identify several issues raised as a result of my adoption of these LA solutions—useful for a general discussion of how LA solutions are conditioned by everyday practice in real settings. In addition to the limitations of some LMS, these issues include the following: the need for multiple metrics and benchmarks for a context-rich understanding of engagement and effective iteration; the difficulty of avoiding university (technical) support for small-scale LA initiatives; the importance of recognizing that ethical grey areas can appear without being anticipated (and be overlooked); the need to accept that some teachers could be ignorant of their university's broader LA initiatives and how this might relate to their own classroom-based teaching goals; and, the importance (and difficulty) of gathering LA data unobtrusively.

Keywords: learning analytics, learning management system, Blackboard, Google, higher education

1 INTRODUCTION

This paper documents my attempt to use learning analytics to answer the following question: are my undergraduate students reading my (blended) unit's online content? "Reading my content" involves, at least, clicking on content, measured with the metric "click through rate" (CTR). More specifically therefore, my question has been: "What is the CTR of my content?" I have reasoned that having this data will help me make educated decisions about changes to content and improvements to my unit.

Solving my problem would seem simple enough, as one benefit of the managerialism underpinning the high rate of adoption of learning management system (LMS) at universities (Beer et. al. 2012), is the collection of student data. Indeed, the collection and analysis of large amounts of data by university systems, such as an LMS, is the purview of learning analytics (LA). As my account demonstrates, however, there are several obstacles to accessing and analyzing learning data on a small-scale at the classroom level.

I encountered these obstacles first-hand, and in what follows I explain my attempts to work around them, developing what I call several “LA hacks.” While these hacks provided me with data that gave me an idea of student engagement with my unit’s content (thus enabling me to optimize the content), they also identify several issues specific to the implementation of LA solutions at the classroom level.

2 CONTEXT: BIG DATA, SMALL DATA, AND ORCHESTRATION

LA research is typically the purview of Big Data, and Big Data is defined as a “cultural, technological and scholarly phenomenon” that rests on the interplay of technology, analysis and the belief that larger data sets offer a higher form of intelligence (Boyd & Crawford, 2012, p. 663).

Typically, the focus of Big Data in LA has been on “analyzing institutional data captured by an LMS and other institutional information systems” (Campbell et. al., 2007) to track student interaction, identify behavior change and enable early identification of “at risk” students (Colvin et. al., 2015). For example, a student profile might be built from a weighted combination of demographics, online engagement data (e.g. LMS activity) combined with an assessment of aptitude (Colvin et. al., 2015).

Retention, risk and attrition are often listed first as the core areas of LA analysis (see for example Colvin et. al. 2015), but research does identify the value of LA data for understanding student engagement with content and improving curriculum design (Dawkins, 2016; Howell et. al., 2017). Involved is a focus on “small data” in LA, specifically at the classroom level. In this context of LA, it is also important to note that the person best placed to decide on relevant student engagement metrics and evaluate the data is the teacher or course designer (Macfadyen & Dawson, 2010).

It has also been noted that LA research understands how student activity and learning is complex, involving a variety of technologies across different spaces (Martinez-Maldonado, 2016). Moreover, a useful perspective on student learning is “orchestration,” which recognizes that classrooms/blended-learning scenarios are variable and complex; educators need to adapt technical resources to enable students to achieve their learning goals; and, technology used in LA environments should be practical, minimalist and flexible—so as to prevent hindering the learning activities (Martinez-Maldonado, 2016).

3 LEARNING ANALYTICS AT AN AUSTRALIAN UNIVERSITY

3.1 The problem

It is accepted in the industry that CTR is a fundamental metric for understanding user engagement with content (Raso, 2016), and so it follows that the CTR of university reading content is a useful metric for understanding student engagement with content.

I encountered several problems in my attempt to measure student engagement with content, and the first was (and is) my ignorance about my university’s own Big Data LA initiatives, and what (if any) initiatives could be directly applied to my research question. The perception that “the provision of information about how learning analytics is being used” is “poor or very poor” was shared by most academics in a recent survey (Rogers et. al., 2015). Not knowing where to start regarding the broader LA initiatives of my university, I instead decided to focus my attention on what was easily available to me: my university's LMS (Blackboard) and its "Course Analytics" data.

My next problem was with Blackboard. This LMS does not enable instructors to accurately measure student engagement with content. Blackboard provides student "views" data on a "Content Item" or "Web Links." A Content Item is a container that holds content for a topic of a unit; for example, learning objectives; online lectures; links to readings; and preparation questions. A Web Link is a link to a page on the internet, and Web Links can only be positioned before or after a Content Item. Data of views of a Content Item or Web Links is problematic for understanding student engagement. Views of a Content Item is only a general metric, and more ideal data would drill down and explain, for example, engagement with readings (CTR). Web Links *do* provide specific engagement data, but their positioning complicates the user-pathway, negatively impacting user engagement. A leading principle of user experience design (UX), defined as the "optimization of a product for effective and enjoyable use" (Lamprecht, 2017), is the "law of pithiness" from Gestalt psychology, which emphasizes that websites that are easy to use and achieve their objectives are clear, ordered and simple (5 psychological principles of high converting websites, n.d.). For example, this can involve less calls to action (CTA), and/or less steps in the conversion funnel—since "one naturally expects fewer users at each step" (Stokes, 2013, p. 504). In terms of measuring student engagement with content on Blackboard, even the most basic application of "pithiness" would involve positioning links, using simple HTML, in a clearly defined pathway within a Content Item. More sophisticated design might place the link beside an image, after a CTA—or even replace the link with a button.

3.2 The hacks

The first hack was my initial attempt to find a way around Blackboard's limitations (noted above) and measure student CTR of unit content in the LMS without disrupting user experience. I was aware of research that has tackled "reading compliance" and, using quizzes and surveys, has found that as little as 20 to 30 percent of students complete weekly readings (Burchfield & Sappington, 2000), but my professional experience led me to CTR as a method for understanding the problem. Also, I wanted to be able to implement the solution in several weeks of the unit so as to enable me to test the CTR of different content formats (for example, text and video and audio) and genres (for example, academic and non-academic text).

I was unsure of the technical capabilities of the LMS, so I contacted the University's Blended Learning Team (BLT) and, after some conversation, the following solution was implemented in two weeks of the unit. I inserted a hyperlink in the weekly Content Item which directed students to an HTML page outside the LMS. On this page, I inserted another hyperlink for downloading a PDF file of the reading content, with the following CTA: "Download the PDF (xMB)." Data from two metrics, page views of the HTML page (via Google Analytics) and downloads of the PDF (via server files), provided me with insights into student engagement with content that week. The data was provided by the BLT since I did not have access to either of the sources. I considered the CTR of the first link to be indicative of students' intention to read the content—in other words, engagement. Also, I was resigned to a decline in CTR between the first and second step of the process given the two-step process complicated the user pathway.

The second hack was another attempt to gather data on student engagement with content without disrupting user experience. This time I also wanted to completely minimize BLT support. Drawing further on my industry experience, this time in mass email optimization, I decided to gather data on student engagement with content by emailing course content to students and tracking the CTR of links in the email. I was aware mass email is typically only opened by a small number of subscribers relative to the number who received the email, and even less subscribers typically click on links (Email marketing benchmarks, n.d.), but I reasoned that comparing CTR to established benchmarks would nevertheless provide me a relative understanding of student engagement with content. My university uses Blackboard to send student emails, and since I was unsure of the capabilities of Blackboard's email I needed to contact the BLT for advice. I was informed that Blackboard does not provide CTR or open rate data from its emails.

Since this was the case I decided to use a third-party email service provider (ESP) that I had some professional experience using: MailChimp. Naturally I considered whether permission was required from my university. The learning and education portfolio had flagged with me—in a separate context—the need for ethics approval for teaching experiments that involved publishing student data. Since I did not intend to publish any data from the email hack, I decided early on not to submit a formal application. I needed to contact the BLT for assistance with transferring student email addresses from the University's database to MailChimp, and at this stage I was advised that, despite the University having no policies/guidelines preventing me from using a third-party ESP, I needed to confirm that MailChimp could provide the following: adequate "support" (if students needed assistance with the emails); appropriate "data storage" of email data; and "data retention" (in case the email data needed be extracted in the future (Saliba, personal communication, July 29 2015). After contacting MailChimp and reassuring the University that MailChimp could provide all the above, it became apparent that I needed to seek permission from the Academic Registrar. I did so, and the experiment was permitted (since data gathered would be de-identified), but I was nevertheless asked to seek ethics approval for privacy and data management reasons. I contacted Human Ethics and was informed of the following conditions of the experiment: I was permitted to discuss de-identified data from the emails for internal use in the University without ethics approval; and I could publish a discussion of the experiment's methodology without ethics approval (Pangilinan, personal communication, October 9 2015).

Eventually, I built two optimized and responsive emails myself (by modifying MailChimp's templates) with embedded links to the readings. Students were not prompted to expect the special emails. The first email was sent twice, on a Friday and Tuesday, and this was based on the recommended send times suggested by MailChimp (Insights from MailChimp's send time optimization strategy, 2014). Since the Friday send received a higher engagement, the second email was only sent on a Friday. After each send I used data from the ESP's campaign report to analyze open rate and CTR of links to unit content. Furthermore, I optimized content placement in the third send based on CTR data from the first and second sends: I placed the content I considered most important in the position that previously had the highest CTR.

I also designed a third LA hack, again to understand student engagement with content. This time I was interested in an LA solution I could completely implement and manage myself, multiple times throughout the unit. I wanted to completely avoid BLT support and having to seek permission from University stakeholders. This meant that a key requirement was for the hack to be an "off-the-shelf" solution that I could embed myself in the HTML of a Content Item in the LMS. And like the other hacks, I did not want to implement technologies that would disrupt the students' reading pathways on the LMS.

I devised the following solution. I used the weekly topic content to design an online activity—for example, a series of short-answer questions that asked students to apply a concept from the readings. I created these activities using Forms on Google Drive and embedded them within the Content Item and strategically placed on a user pathway. And, I asked students to identify themselves and their class day and time in a separate field of the form. I reasoned that permission from the University was not required since I did not intend to publish any of the data gathered in the forms (there would be no data identification issue), and since access to the Google Forms was only through the University's password-protected LMS (there was no privacy issue).

I asked students to complete the activities before attending class. I planned to use the students' answers to evaluate their understanding of the reading and before-class preparation; but also, it would be clear when there was no response in the forms who had not attempted the reading at all. I could have implemented a similar activity using a "Discussion Board" in Blackboard, but I choose to use Google because I could position the Forms anywhere in the Content Item. In addition, I could see a summary of responses at a glance with Google's analytics dashboard.

I implemented the third hack three times during the thirteen-week unit. I analyzed the response-rate for each activity relative to the following: the week of the topic (i.e. was it early in the unit, in the middle, at the end; and, was it before an assessment, after an assessment, or in the intra-session break?); the number and complexity of online lectures and readings; the activity's position relative to other content in the Content Item; and finally, in terms of the number and complexity of the questions themselves.

3.3 Results/Discussion

On the one hand, I can argue that the hacks were successful. This is because they each provided me with LA data I could not otherwise have obtained. Moreover, evidence that demonstrates the value of solutions designed to track student engagement with content is an increase in student satisfaction in feedback surveys from 2016 (n=58). In the units reviewed, I implemented several "analytics hacks," including the third initiative noted above (Google Forms). In one of the surveys, all questions in the survey show an increase in student satisfaction since 2015, but the largest effect can be seen in the improvement of scores for "Learning design" (average score of 3.5 [2015] vs. 4.4 [2016]) and "Learning resources" (3.3 [2015] vs. 4.2 [2016]).

On the other hand, evaluating the hacks is more complex and warrants further discussion. What were the critical issues and lessons learned from the adoption of these LA solutions at the classroom level? Orchestration is a useful perspective for unpacking this complexity. Orchestration understands that student activity and learning involves multiple stakeholders and a variety of technologies distributed across different spaces. In addition, an orchestration perspective emphasizes that LA technologies should be practical, minimalist and flexible (Dillenbourg, 2013 in Martinez-Maldonado, 2016); and, an orchestration framework involves a notion of iteration; for example, consider the "four stage iterative process": teachers access data; assess data; develop insights from data; and, introduce new insights (Verbert, 2013 in Martinez-Maldonado, 2016, p. 71).

The perspective of orchestration reveals the limitations of some insights offered by the hacks. Each hack was implemented in isolation in a week of a unit (or in different units). As noted in previous LA research, student learning typically involves a variety of tools across a variety of spaces. To better attend to the way students learn, data gathered from a combination of sources could more accurately describe student engagement with content. This said, future research might implement all hacks simultaneously in one week. Of course, using multiple metrics to measure engagement is best practice in website optimization (Patel, 2016), and while this approach is noted here to emphasize what more comprehensive LA solutions could involve, it also needs to be said that a teacher's workload may not allow for such a detailed, multi-faceted approach in a single week of a unit—and this was certainly the case for me.

Another issue raised by the hacks is the complexity of stakeholder involvement. First, BLT assistance was necessary for the first two hacks, despite my concerted effort to design and implement them myself. This suggests that technical support may be unavoidable for teachers interested in implementing LA solutions, and this may be for the simple reason that university BLTs have exclusive access to data sources. The level of technical support needed can impact on the viability of the LA solutions, and in the case of the first hack I decided not to repeat the hack in other units precisely because the technical support needed added a significant layer of complexity. Second, my dialogue with university stakeholders proved there was some uncertainty about the permissions required for the email hack. Perhaps this "uncertainty" justifies my concern that university approval processes could have a chilling effect on future small-scale LA experiments—noted as crucial for driving innovation in today's technological teaching space (Office for Learning and Teaching, 2015, p. 38).

Human Ethics, of course, is a necessary stakeholder. Ethics in LA is an important area of research and there is not scope in the current study to engage with this issue in detail. The complexity added by approval processes, such as ethics, to small-scale LA projects needs to be noted; but also, discussion of ethics approval processes raises another issue relevant here: ethics grey areas in research and the question of when ethics approval is necessary. The third hack illustrates one such grey area. I reasoned that ethics approval was not necessary for my implementation of Google Forms in several weeks of the unit, but I realize now that future iterations of this hack should consider the (potential) privacy implications regarding students' sharing of information on a collaborative Google Form, as well as (potential) data-retention issues specific to Google Drive. This hack identifies how ethical grey areas can appear without being anticipated—and can easily be overlooked—in rapidly advancing technological teaching spaces.

The teacher is another crucial stakeholder in LA orchestration, and my role in the above hacks identifies several important issues. The biggest issue was my ignorance about the LA initiatives operating at my university and their potential relevance for my own teaching practice. Related is my competency, and confidence, with LA technologies—and teachers' concerns with the technical knowledge required for LA solutions has been noted in previous research (Rogers et. al., 2015). In my case, I have some basic expertise from previous professional work that has helped me design and implement the hacks noted above, and recognize their limitations. But this may be a level of technical expertise missing for many teaching academics at university.

In terms of orchestration's emphasis on practical, minimalist and flexible technology, the hacks were successful in so far as they were easily embedded *within* the Content Items. But did these hacks enable a positive user experience? It needs to be acknowledged that in terms of the first hack, any positive effect for user experience of embedding the link in the Content Item was probably cancelled by the two-step solution implemented. In addition, this hack required that content was PDF format, which limits the resources that can be used and has been noted in industry research as a disliked format since users would rather not download content (Nielsen, 2001). In addition, the second hack was hampered by characteristically low open rates and CTR of mass email (Email marketing benchmarks, n.d.), and the third hack was another example of a two-step process where students were taken out of the LMS. In sum, I acknowledge how important it is to implement "minimalist" LA solutions that do not disrupt the user experience, but these examples demonstrate how difficult this can be to achieve.

In terms of iteration in orchestration, minimal changes were made to content because of the hacks, making the usefulness of the hacks limited. One reason iteration was limited is the lack of benchmark data against which to measure engagement. While email benchmark data exists for open rate and CTR (Email marketing benchmarks, n.d.), I am unaware of benchmark data for CTR in an LMS, or engagement in an online activity using Google Forms. As a result, the optimization noted above was largely guesswork. Another reason iteration was limited is a direct result of the design of university units themselves. In an ideal scenario, a teacher would change content based on engagement data; but it is typically the case that course content is decided months in advance, approved by directors of academic programs, and is unable to be changed "on the fly." In any case, it is unreasonable to notify students about changes to content less than a week in advance, which also makes last minute changes based on CTR data unrealistic. It is clear, therefore, that the iterative adoption of LA solutions can be difficult to effectively achieve in a university classroom context.

Finally, a critique of the hacks reveals the importance in classroom-based adoption of LA technologies of understanding the context of data (Boyd & Crawford, 2012). It needs to be noted that CTR is not necessarily reflective of reading since many internet users, students included, click—and even share—content without reading it. Also, CTR can vary according to a user's opinion of where the link is taking them, as well as the wording of the link text, and the placement of the text. Put simply, each of these factors constitute the context of the data and need to be considered and accounted for when designing and evaluating LA solutions.

4 CONCLUSIONS

This study documents my implementation of LA classroom "hacks," designed to provide me with data (student engagement with content) I could not obtain from my LMS. I reasoned that accessing this data could help me better understand student engagement and improve the design and content of my units.

The hacks themselves illustrate a range of ways engagement data can be gathered—using free, "off-the-shelf" third-party technologies. Importantly too, the processes involved in designing and implementing the hacks identify key issues in the orchestration of LA solutions at the classroom level. Most salient are: the need for multiple metrics and benchmarks for a context-rich understanding of engagement and effective iteration; the difficulty of avoiding university (technical) support for small-scale LA initiatives; the importance of recognizing that ethical grey areas can appear without being anticipated, and be overlooked; the need to accept that some teachers could be ignorant of their university's broader LA initiatives and how this might relate to their own classroom-based teaching goals; and, the importance (and difficulty) of gathering LA data unobtrusively.

There are clearly issues involved in the adoption of LA and orchestration solutions at the classroom level and these need to be resolved, but we should nevertheless bear in mind the opportunities afforded by the internet and its data traces, for us (teachers) to roll up our sleeves and cobble together another perspective on our students.

REFERENCES

5 psychological principles of high converting websites (n.d.). Retrieved 2018, December 7 from https://blog.kissmetrics.com/psychological-principles-converting-website/

Beer, C., Jones, D. & Clark, D. (2012). Analytics and complexity: Learning and leading for the future. In M. Brown, M. Hartnett & T. Stewart (Eds.), *Future challenges, sustainable futures*. In Proceedings ascilite Wellington. 78-87

Boyd, D. & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Societ*y, 15(5), 662-679, DOI: 10.1080/1369118X.2012.678878

Burchfield, C. M. & Sappington, J. (2000). Compliance with required reading assignments. *Teaching of Psychology*, 27(1), 58-60.

Campbell, J.P., DeBlois, P. B & Oblinger, D.G. (2007). Academic analytics: A new tool for a new era. *Educause Review*. 42(4).

Colvin, C., Rogers, T., Wade, A., Dawson, S., Gasˇevic´, D., Buckingham Shum, S., & Fisher, J. (2015). Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement. Sydney: Australian Office for Learning and Teaching. Retrieved from http://www. research.ed.ac.uk/portal/files/21121591/Final_Report_190615.pdf

Dawkins, R. (2016). Content strategy: a lesson from the industry for university learning analytics. In S. Barker, S. Dawson, A. Pardo, & C. Colvin (Eds.), *Show Me the Learning*. Proceedings ascilite Adelaide. 172-181.

Email marketing benchmarks (n.d.) Retrieved 2018, December 7 from https://mailchimp.com/resources/research/email-marketing-benchmarks/

Howell, J. A., Roberts, L. D., Seaman, K. & Gibson, D. C. (2017). Are we on our way to becoming a "helicopter university"? Academics' views on learning analytics. *Tech Know Learn*. DOI 10.1007/s10758-017-9329-9

Lamprecht, E. (2017, January 13). The difference between UX and UI design—A layman's guide. Career Foundry. Retrieved from https://www.careerfoundry.com/en/blog/ux-design/the-difference-between-ux-and-ui-design-a-laymans-guide/

Martinez-Maldonado, R. (2016, January). Seeing learning analytics tools as orchestration technologies: Towards supporting learning activities across physical and digital spaces. *CEUR Workshop Proceedings: Proceedings of the First International Workshop on Learning Analytics Across Physical and Digital Spaces co-located with 6th International Conference on Learning Analytics & Knowledge* (LAK 2016), 2016, 1601 (pp. 70 – 73).

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an ''early warning system'' for educators: A proof of concept. *Computers & Education*, 54(2), 588–599. http://dx.doi.org/10.1016/j.compedu.2009.09.008.

Neilsen, J. (2001, June 10). Avoid PDF for on-screen reading. *Nielsen Norman Group*. Retrieved from https://www.nngroup.com/articles/avoid-pdf-for-on-screen-reading/

Office for Learning and Teaching. (2015). Student retention and learning analytics: A snapshot of Australian Practices and a framework for further advancement. Retrieved from http://he-analytics.com/wp-content/uploads/SP13-3249_-Master17Aug2015-web.pdf

Patel, N. (2016, May 18). How to measure reader engagement and loyalty using Google Analytics. *Neilpatel.com*. Retrieved from http://neilpatel.com/2016/05/18/how-to-measure-reader-engagement-and-loyalty-using-google-analytics/

Raso, A. (2016, March 28). How to measure engagement the right way. *Content Marketing Institute*. Retrieved from http://contentmarketinginstitute.com/2016/03/measure-engagement-right/

Rogers, R., Colvin, C., West, D. & Dawson, S. Learning analytics in Oz: What's happening now, what's planned, and where could it (and should it) go? LAK '15 *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*. (pp. 432-433).

Stokes, R. (2013). *eMarketing: The essential guide to marketing in a digital world* (5th ed.). Quirk education. Available: www.quirk.biz/emarketingtextbook

Straumsheim, C. (2016, March 2). Read and unread. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2016/03/02/study-explores-impact-social-media-texting-email-use

LessonSpectrum: Visualizing Teaching and Learning Activities in English Conversation Lessons

Zilu Liang
Faculty of Engineering, The University of Tokyo, Japan
z.liang@cnl.t.u-tokyo.ac.jp

Satoshi Nishimura, Takuichi Nishimura
Artificial Intelligence Research Center, Japan
National Institute of Advanced Industrial Science and Technology
{satoshi.nishimura, takuichi.nishimura}@aist.go.jp

Mario Alberto Chapa-Martell
CAC Corporation, Japan
Mchapam0300@gmail.com

ABSTRACT: Understanding teaching and learning activities in a lesson has multiple benefits for teachers, students and educational institutes of second language education. In this paper, we proposed a data analysis tool named LessonSpectrum that automatically generates visualization of classroom activities in the form of heat maps. We presented the interaction among multiple stakeholders in the development of LessonSpectrum as well as the results of preliminary adoption of this tool.

Keywords: Classroom activity, data visualization, classroom orchestration, English education, Javascript, D3.js.

1 BACKGROUND

In response to the need of improving English communication skills of engineering students, the Global Center of Innovative Engineering Education (GCIEE) of the University of Tokyo started the Special English Lesson (SEL) program in 2005. Seven English language schools dispatch native speaking English teachers to give lessons on the campus of the university. Students decide on which lesson to take, and they directly register to the language schools. The responsibility of the GCIEE is to coordinate between students and the language schools and to ensure the quality of lessons. As such there are three stakeholders in the SEL program and the needs of each stakeholder are summarized below: the university students who want to maximize their English learning outcomes through SEL classes; the GCIEE that needs to help students to select the classes that best suits their purposes and objectives and to monitor the quality of teaching; and the English teachers who want to know how they can improve their teaching (such as time management).

Teaching assistants (TA) observe the lessons and give comments in notebooks. Although 12 years of TA notes has been collected, these data are largely descriptive, ad hoc, and provide little information of what actually happened in a lesson. These data has generated no insights as analysing these data requires large amount of time and simply became unfeasible. As a result, it has been beyond the capacity of GCIEE to guide the students to select the most suitable lessons and to give actionable feedback to the teachers.

To address the above problem, the authors helped the GCIEE develop a data analysis tool named LessonSpectrum which automatically generates visualization of the sequence of classroom activities in the form of a heat map. The visualization provides rich information on teachers' real-time management of multiple classroom activities and numerous teaching actions and enables comparison between different lessons. In what follows we present the development of LessonSpectrum and the results of preliminary adoption of this tool by the educational institute and the teachers.

2 TOOL DEVELOPMENT

LessonSpectrum is a Javascript application that renders visualization of the sequence of classroom activities in a web browser. Users simply need to upload a csv file that comprises the records of the classroom activities following the coding scheme described in Section 2.2, LessonSpecturm automatically creates a grayscale heat maps of learning and teaching activities.

2.1 Development of Classroom Activity List

The activities occurred in second language classes are highly repetitive and share common features across classes. Compiling a list of common classroom activities enables us to analyze the structures of different classes under the same framework. To create such a list, the SEL program coordinator observed six classes in full length and recorded all the activities occurred in those classes using a mobile application named aTimeLogger (www.atimelogger.come). The recorded data comprised the time stamp and duration of each activity as well as observer's comments. These data were exported into CSV files from aTimeLogger. Developers then read through all the recorded class logs and extracted a list of initial classroom activity codes such as "having group discussions", "giving a presentation", "explaining new words and phrases" and "checking answers with the whole class". These codes were broadly categorized into teaching and learning activities. The SEL program coordinator and an English teacher revised the activity list to ensure that the items are common activities in English conversation classes.

2.2 Coding of Classroom Activities

After compiling the lists of teaching and learning activities, the SEL program coordinator and an English teacher separately ranked the activities according to perceived educational benefits. Afterwards they compared their ranking together and resolved conflicts through discussions. The outcomes were two lists of teaching and learning activities ordered by education benefits. These lists were then passed to developers to map to numerical color codes following the color scheme of the Javascript library D3.js. In this scheme, a color code of 0 corresponds to black, while that of 1 corresponds to white. Table 1 summarizes the code lists of teaching activity and learning activities. Note that this list is not exhaustive as it is limited to the teaching procedures in the observed classes. This list should be expanded if new class activities are observed in the future.

Table 1: Activity coding list.

Teaching Activity	Color Code	Learning Activity	Color Code
Challenge Students' Opinions	0	Debate with Teacher	0
Have Impromptu Discussions with Students	5	Have Group Discussions	5
Ask Follow-up Questions	10	Have Impromptu Discussions with Teacher	10
Give Learning Tips	15	Play a Game	13
Give Comments	20	Ask Follow-up Questions to Presenter	15
Share Personal Experience	25	Give Group Presentation	20
Answer Questions from Students	30	Do Role Play	21
Correct Mistakes	35	3-2-1	22
Explain Grammar Points	40	Wrap Up Group Discussions	25
Explain New Words and Phrases	45	Compare Answers in Group	30
Coordinate Whole Class Discussion	50	Volunteer to Answer Questions	35
Observe Students With Verbal Interaction	55	Brainstorm on Discussion Topics	40
Observe Students with Occasional Verbal Interaction	60	Answer Teacher's Questions when Called	45
Observe Students with No Verbal Interaction; Take Notes	62	Listen to Audio and Answer Questions	50
Observe Students with No Verbal Interaction	65	Do a Quiz or Exercise	55
Compare Answers with Whole Class	70	Ask Questions to Teacher	60
Give Pronunciation Drill	75	Repeat After Teacher	65
Give Listening Practice	80	Read Out	70
Explain Tasks and Take Questions	85	Listen to Teacher and Take Notes	75
Listen to Students	90	Listen to Teacher	80
Write on White Board	98	Stand-by	99
Assign Homework	98	Have a Break	100
Check Teaching Plan	98		
Check Attendance	98		
Have a Break	100		

2.3 Post-hoc Analysis

Many features can be extracted from the heat maps for quantitative evaluation of class. For example, the mean and standard deviation of the color codes may be used to estimate the educational impact and the diversity of a class. Lower mean value indicates more educational benefits while larger deviation corresponds to a class with more kinds of activities.

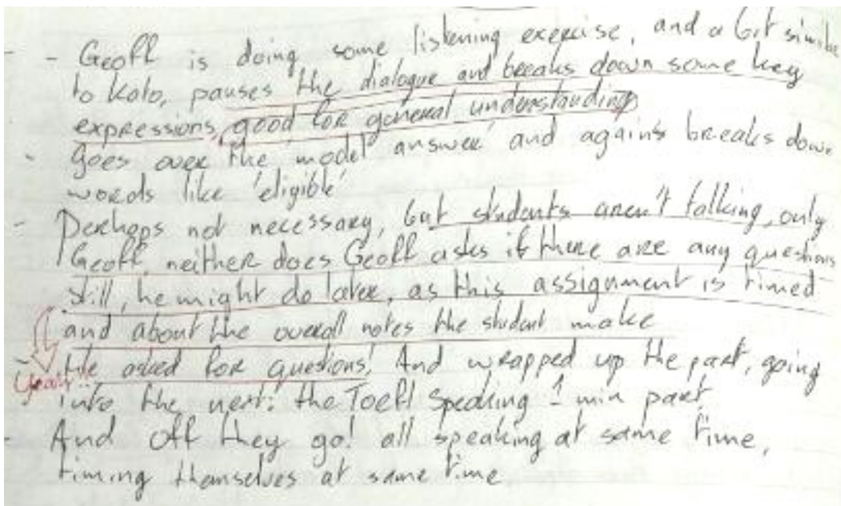
	A	B	C	D	E	F
1	TimeStamp	Minutes	LearningCode	LearningActivity	TeachingCode	TeachingActivity
2	19:01:00	1	80	Listen to the teacher	85	Explain tasks and took questions
3	19:02:00	2	5	Have group discussion	100	Check teaching plan
4	19:03:00	3	5	Have group discussion	100	Check teaching plan
5	19:04:00	4	5	Have group discussion	100	Check teaching plan
6	19:05:00	5	45	Answer teacher's question when called	10	Ask follow up questions
7	19:06:00	6	45	Answer teacher's question when called	10	Ask follow up questions
8	19:07:00	7	45	Answer teacher's question when called	10	Ask follow up questions
9	19:08:00	8	45	Answer teacher's question when called	10	Ask follow up questions
10	19:09:00	9	45	Answer teacher's question when called	10	Ask follow up questions
11	19:10:00	10	45	Answer teacher's question when called	10	Ask follow up questions
12	19:11:00	11	45	Answer teacher's question when called	10	Ask follow up questions
13	19:12:00	12	45	Answer teacher's question when called	10	Ask follow up questions
14	19:13:00	13	45	Answer teacher's question when called	10	Ask follow up questions
15	19:14:00	14	45	Answer teacher's question when called	10	Ask follow up questions
16	19:15:00	15	45	Answer teacher's question when called	10	Ask follow up questions
17	19:16:00	16	45	Answer teacher's question when called	10	Ask follow up questions
18	19:17:00	17	45	Answer teacher's question when called	10	Ask follow up questions
19	19:18:00	18	80	Listen to the teacher	85	Explain tasks and took questions
20	19:19:00	19	45	Answer teacher's question when called	70	Compare answers with the whole class
21	19:20:00	20	45	Answer teacher's question when called	70	Compare answers with the whole class
22	19:21:00	21	45	Answer teacher's question when called	70	Compare answers with the whole class
23	19:22:00	22	45	Answer teacher's question when called	70	Compare answers with the whole class
24	19:23:00	23	45	Answer teacher's question when called	70	Compare answers with the whole class
25	19:24:00	24	45	Answer teacher's question when called	70	Compare answers with the whole class
26	19:25:00	25	45	Answer teacher's question when called	70	Compare answers with the whole class

Figure 1: An example of coded data.

3 TOOL ADOPTION

3.1 Adoption by Educational Institute

The GCIEE has been relying on TA notes for gaining information on classroom activities in the SEL program. Last year, the GCIEE attempted to replace paper-based TA notes with digital word files to reduce the hurdle of post-hoc analysis on the data collected. Since teaching assistants only spent approximately 15 minutes in each class, however, the TA notes only consists snapshots of a class. Therefore, analysis on these data only produce fragmented, if any, information. In contrast, LessonSpectrum renders visualization of the full spectrum of activities occurred in a class. Figure 2 demonstrates the comparison of paper-based TA notes, digital TA notes, and LessonSpectrum. For GCIEE, the cost involved in developing LessonSpectrum is mainly the time and human resources required to compile the activity coding lists through recording classroom activities in many classes and the technical implementation of the tool. On the other hand, LessonSpectrum will bring multi-fold benefits to GCIEE, including better matching between students and classes, increased satisfaction of students, and actionable feedback for teachers. Privacy and copyright issues were initially the local constraints for the adoption of the tool. Since the visualization does not expose any sensitive, identifiable, and copyrighted data such as pictures or videos, these constraints were eventually released, and the SEL program coordinator decided to integrate LessonSpectrum into the SEL echo-system. Though LessonSpectrum has not yet been tested with students, the feedback from SEL teaching assistants were positive: "*These graphs can be useful for students when they need to choose from a pool of teachers. Maybe some students will like to take very dynamic classes*". In the coming semester, the class spectrum of different teachers will be demonstrated to students to help them select the most desired class. In the meanwhile, the SEL program coordinator will continue collecting activity logs from more classes to enrich and enhance the activity coding lists.



- Geoff is doing some listening exercise, and a bit similar to Kolo, pauses the dialogue and breaks down some key expressions, good for general understanding
- goes over the model answer and agains breaks down words like 'eligible'
- Perhaps not necessary, but students aren't talking, only Geoff, neither does Geoff asks if there are any questions still, he might do later, as this assignment is timed and about the overall notes the student make
- He asked for questions! And wrapped up the part, going into the next: the Toefl Speaking 1 min part

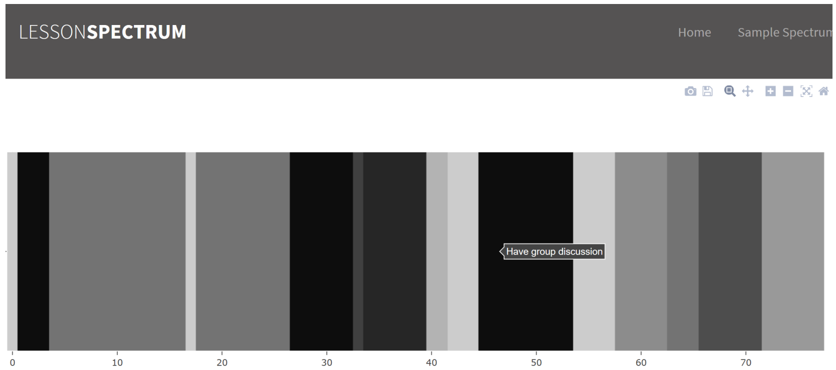
Great

- And off they go! all speaking at same time, timing themselves at same time

(a) Paper-based TA notes

What did you see in the class (e.g., what were the students and the teacher doing, on what topic were they discussing, what materials were they use, what tasks did the teacher give to students, how did the teacher help/correct the mistakes of the students, how was the atmosphere, what assignment/homework did the teacher give to students, etc.) **The teacher was explaining about video games, binge watching, dopamine and other related vocabulary. The teacher then moved to other Vocabulary words like Game Theory etc.**	
How do you think about the teacher's performance today (e.g., attitude, energy, motivation, etc.) **The teacher was very patient in answering the questions. He was explaining differences between corroborate, collaborate and cooperate.**	*How do you think about the students' performance today (e.g., reaction, attitude, energy, motivation, etc.)* **The students were asking very intriguing questions while learning. They were also going for extra vocabulary other than the reference material.**
What were the problems you noticed in today's class?	
What were the good things you noticed in today's class? **The teacher was explaining them a new concept. He explained them bout Tetris effect.**	
Is there anything else you want to share with us?	

(b) Digital TA notes



(c) LessonSpectrum

Figure 2: Comparison of current practice of educational institute GCIEE (a) paper-based TA note, (b) digital TA note, and LA-enhanced practice (c) LessonSpectrum.

3.2 Adoption by Teachers

Most of the teachers dispatched from the language schools are well-prepared for the classes. Some of them always write down the tasks or the schedule of the class on black board based on their lesson plan at the beginning of the class. However, currently there was no way for teachers to figure out how the lesson plan turned out especially in terms of time management. We have noticed that time management was an obviously problem for novice teachers, which in turn affect the satisfaction of students. The LessonSpectrum will enable teachers to examine their real-time management of multiple classroom activities and teaching actions as well as make it possible to compare different classes quantitatively. If used over long term, it will also help teachers discover changes in their teaching practices. In addition to the multiple benefits of LessonSpectrum for teachers, flat learning curve and low cost of installation highlight its potential for large-scale adoption.

In the process of developing LessonSpectrum, we collected feedback from two English teachers. They mentioned that they obtained useful information from the lesson spectrums, including the activities occurred in a lesson, the pace of the lesson, and teachers' teaching style. They believed that these plots have potential applications. "*The visualization can be used for evaluating teaching methods*." (Teacher A)

In addition, both teachers have positive attitude towards adopting LessonSpectrum in their teaching because they believed that such visualization can help them reflect on their teaching. "*It is always helpful to know whether a class was delivered as it was planned; in this way I can improve my teaching*." (Teacher A) "*Sometimes one's perception of our own activities might not be accurate. I may have the idea that I delivered an active lecture but then it turned out that I spent too much time doing just a few activities. Having these graphs will provide objective feedback*." (Teacher B)

However, there are two obstacles for seamless integration of LessonSpectrum into teaching routines. First, teachers need to record the classroom activates on their own, which may inevitably interrupt their lesson plan. Addressing this constraint requires developing intuitive user interface for the tool to reduce the burden of manual log. Second, the policies and restrictions of each language school may require permission of the authorities before teachers can use the tool in their classes. Therefore, the SEL program coordinator planned to introduce this tool to the representatives of each language school in the annual review meetings to be held in March 2018.

4 FUTURE WORK

In the next step, we plan to develop a mobile application based on the concept of LessonSpectrum to facilitate large-scale adoption. We also plan to investigate the impact and benefits of such adoption.

REFERENCES

Wilkinson, L., and Friendly, M. (2009). The History of the Cluster Heat Map. *The American Statistician 63*(2): 179–184. http://www.tandfonline.com/doi/abs/10.1198/tas.2009.0033

Towards a Comprehensive Learning Analytics Methodology for a Real Educational Platform: A Case Study

Luis de-la-Fuente-Valentín
Universidad Internacional de La Rioja (UNIR), Spain
Luis.delafuente@unir.net

Jingjing Zhang
Beijing Normal University (BNU), China
Jingjing.zhang@bnu.edu.cn

Iván Arakistain
Universidad Internacional de La Rioja (UNIR), Spain
Ivan.arakistain@unir.net

Qinhua Zheng
Beijing Normal University (BNU), China
Zhengqinhua@bnu.edu.cn

Daniel Burgos*
Universidad Internacional de La Rioja (UNIR), Spain
Daniel.burgos@unir.net
**Corresponding author*

ABSTRACT: There exist many approaches that enable learning analytics within the educational framework, but their adoption is hindered by the capability of integrating teachers' inquiries into analysis and by the lack of flexibility to adapt analytics to meet teacher specific needs. This study investigates how to use existing tools to build a framework able to respond teachers' inquiries when they are posed in natural language, offering the answers from a data-driven perspective. As a result, this paper proposes a framework to integrate teachers' inquiries into the analysis of a massive dataset using the IBM Watson Analytics tool and presents a case study that validates that framework. The massive dataset contains over 10 years' student behavior data stored in a LMS of a Chinese university in SQL format with a size of over 1 Giga bytes. The results show that analytics can help to answer generic questions that can be formulated using natural language. We also analyze the impact of the data curation process on the quality of the obtained answers.

Keywords: natural language, teacher questions, data-driven answers, *IBM Watson Analytics*

1 INTRODUCTION

The research field of Learning Analytics has generated a great number of on the methods and tools that allow measuring, collecting, analyzing and reporting of data about learners and their contexts (Baker & Inventado, 2014). However, specific barriers can limit the wide adoption of learning analytics tools. A. Wilson et al. analyzed four potentially problematic aspects: the inconclusiveness of empirical studies;

somewhat simplistic conceptions of learning analytics data and methods as part of some generic species, Big Data; choices about data, algorithms and interpretation; and issues around disciplinary and finer-grained differences in pedagogical and learning approaches (Wilson et al., 2017).

Learning Analytics methods and tools can be built from a generic or case specific perspective. The former is (claimed to be) able to support a wide range of pedagogical methods. Some well-known Learning Analytics software tools are SNAPP or LOCO-Analyst (Dawson et al., 2010; Jovanovic et al. 2008). However, teachers' low data literacy competences for analyzing and interpreting complex educational data can limit the access to analytics (Marsh & Farrell, 2015). On the other hand, case specific analytic tools are devoted to provide answers to very specific questions (i.e. (Gasevic et al., 2017) presents a study to detect learning strategies) while they are difficult to adapt to other scenarios.

In such a landscape, this study poses the following question: Can natural language query capabilities improve teacher´s capabilities to understand student´s performance? To the best of our knowledge, there is no framework that enable such an experiment and therefore the actual research question of this work is one step behind: can we build with existing tools a framework that understands natural language questions and provide meaningful answers for the teaching/learning scenario?

This study proposes a pedagogically neutral learning analytics framework able to answer natural language queries formulated by teachers. The resulting framework will be able to provide answers formulated from any pedagogical perspective.

This paper is structured as follows: Section 2 describes the applied methodolody and the case study where they were deployed. Section 3 analyzes how the analytic work was carried out in order to respond to the teacher inquiries and followed by the discussions of its implications. In Section 4 conclusions are drawn and guidelines for future work are outlined.

2 METHODOLOGY

2.1 Workflow

Step 1. Data collection: during this step, a computer-assisted system collects educational data in one or several databases. Strictly speaking, this step is not part of the framework since the educational platform is decoupled of the analytical framework.

Step 2. Data transformation: this framework use IBM Watson Analytics to analyze data and therefore the dataset must be translated from the original format to a Watson Compatible one

Step 3. Data importation: It is necessary to feed the analytics software with data. Once data is imported, IBM Watson Analytics calculates a data quality index, representing its readability for the tool. As a result, a data curation process may be required to improve this index.

Step 4. Data inquiry: during this step, teachers identify specific educational aspects that they wish to investigate in order to improve student´s performance. This last step is executed by the teacher (the actual educational practitioner) while the previous steps should be executed by technical experts.

2.2 Case study description

As a prototype version for the proposed framework, the proposed steps have been deployed in a case study with authentic educational data. This source dataset contains 100.000 records with user activities that are extracted from Learning Management Systems. The data were neither structured nor normalized, the dataset consisted of a number of tables largely disconnected. It contains over 10 years' student behavior data (between 2005 and 2015) stored in a LMS of a Chinese university. Two Navicat database dumps, named as database_dump_01.sql and database_dump_02.sql account to file sizes of 173MBytes and 907MBytes respectively.

To perform data transformation, the datasets were first imported to SQL server and then exported in CSV format. One noticeable drawback of this way of work is the impossibility of the CSV format to preserve foreign keys (and therefore relationships) among tables. Without relationships, important parts of the information are lost.

Once imported in IBM Watson Analytics, the tool creates a data quality report, which includes an overall average data quality score. This data quality score, indicates the readiness of the data to be analyzed and does not necessarily indicate whether Watson Analytics will provide good predictive or explorative results. In other words, a low data quality score just indicates that data is not suitable for analysis, but Watson Analytics might still provide useful insights and answers about this data. A data curation process is then required to optimize the value of the data quality index. After that, the framework is ready to accept natural language questions.

2.3 Evaluation methods

To assess the framework, some sample questions were collected and posed to the system. The researchers evaluated whether or not the output contains proper data to actually be considered a valid answer to the proposed question.

TQ1. Homework participation. Is there a relationship between doing homework and participating in forums and other activities? A teacher that wants to encourage forum discussions may want to understand why his students use forums, and therefore pose this question as a way to find a way to trigger participation in forums.

TQ2. Homework and final result. Is there a relationship between sending homework early and obtained score? A teacher worried for the homework overload might want to be informed about the homework effects.

TQ3. Platform login and stay time. What drives number of logins and stay time? Teachers could be interested in understanding how to motivate students to access the course material.

3 DISCUSSION

3.1 Results to questions

TQ1. Homework participation. Is there a relationship between doing homework and participating in forums and other activities?

In order to answer this question, the educator could directly type it into IBM Watson Analytics text box. For instance: "*What is the relationship between Homework_NUM and Forum_NUM?*" (that is, the teacher should know column names).

The analytical tool will interpret the question and provide several starting points. Usually, it will propose up to 6 options in the first screen. We found out that IBM Watson Analytics interprets the question correctly, as one of the first proposals is "*How does Homework_NUM compare by Forum_NUM?*".

The answer is provided in a graphical format: the system selects the data visualization solution that best matches with the received question. Figure 1 shows the provided answer for this question, where the teacher will find out that there is certainly a relationship between doing homework and the number of participation in forums.

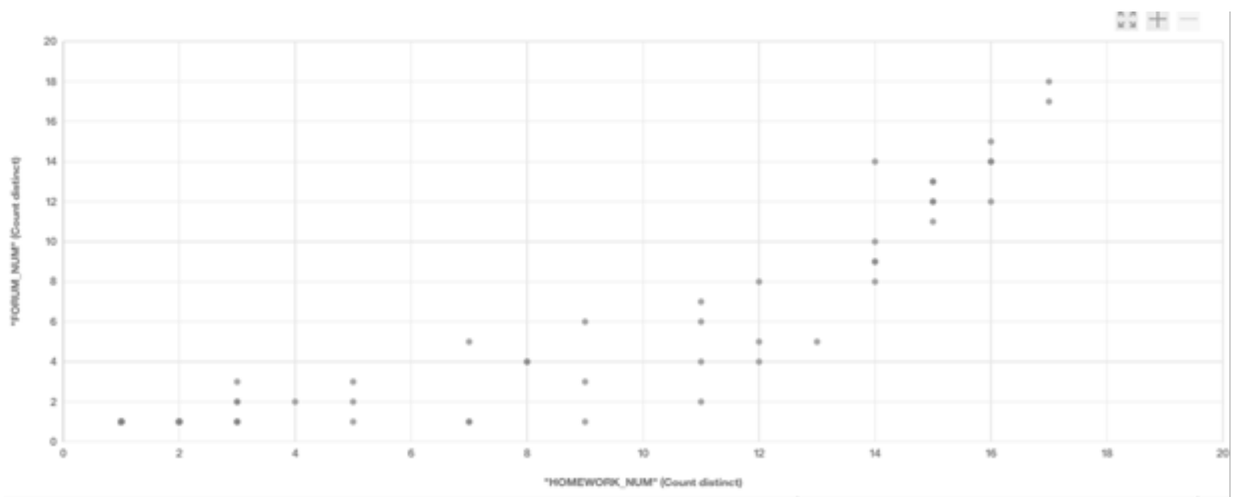


Figure 1 : Relationship between doing homework and forum participation

TQ2. Homework and final result. Is there a relationship between sending homework early and obtained score?

There are 183,000 data registers containing user, course, homework submission date and score information. After obtaining suitable .CSV files with data, rows with "-1" homework score have been deleted as it has been understood that these were not valid scores. Homework scores range between 0 and 100.

This previous data preparation procedure means that quality of data matters, and abnormal values need to be curated in order to obtain meaningful analytics. Once data is curated, IBM Watson Analytics will propose among the answers to the question: "*What is the trend of Homework_Score over Homework_Date?*". The answer given by Watson Analytics, also shaped as a data visualization, relates submission date (aggregated by month) and obtained score in a linear graphic.

TQ3. Platform login and stay time. What drives number of logins and stay time?

This question is easily understandable by IBM Watson Analytics. The phrase structure "*what drives…?*" is well-known by the tool. As a result, the teacher could find out in the proposed responses that Number of login_times are driven by experience, while stay_time is driven by article_num and login_times.

Furthermore, the question can be formulated in another way: "*How can we predict login_time?*" The answer given by the tool is the main element in a decision tree for that purpose and that its prediction strength would be of 66%.

3.2 Adoption

One lesson learned by the researchers of this study while deploying case studies with other learning analytic tools is that the teachers demand tools that provide answers to their questions, and these questions arise spontaneously in the actual teaching situation. Another important lesson learned is that the more understandable the concepts behind the framework, the more prone are teachers to participate in the design process.

Therefore, the main goal of this study is to demonstrate the technical feasibility of a framework able to answer such type of questions. That is, this work is a proof of concept required to engage teachers in a more comprehensive design of the framework. It is the belief of the researchers that this methodology will result in a softer adoption process.

3.3 Stakeholders

As stated in Section 3.2, this study is on step behind the actual development of the tool. Taking this idea into account, the role of the different stakeholders was the following:

- Researchers: responsible of the main design principles of the framework. They applied the lessons learned from other studies and composed a solution proposal. Their goal is to use this proof of concept to engage teachers in later steps of the design, development and validation of the tool. In addition, the researchers designed the validation questions for the prototype. The researchers recognize that teachers might have posed these questions, but they also recognize that this work is one-step behind this phase.

- Developers: the main difficulties found in the presented study have been technical ones (e.g. the need to translate the dataset to a IBM Watson compatible), so the developers have been key to determine whether the technology is ready to provide such a framework. Their output was quite valuable to understand what functionalities can be offered to teachers for the later design.

- Teachers: they will play their role in future steps. The prototype presented in this paper demonstrates the potential of the framework and the teachers should take part in the design and, more important, validation of the actual framework.

4 CONCLUSIONS

In this research work, we have developed a framework to be able to answer generic questions with IBM Watson Analytics by using data from a SQL database. It has been shown that analytics can help to answer generic questions that can be formulated using natural language

While this methodological proposal is strong enough to apply a posteriori analytical approach, the usability of the framework is not ready to be deployed in an authentic scenario and used by non-technical users. Therefore, this case study was devoted to understand the curation process, identify deficiencies and extract lessons learned useful for future implementations.

The question driven analysis has been executed over a curated dataset. One interesting question to analyze is the impact of the data curation process on the quality of the obtained answers. In order to face such analysis, future work will include the elaboration of three different versions of the dataset, stored in separate files. The goal is to easily compare the results of the questions when posed to different versions of the data.

Real adoption of Learning Analytics in the community cluster of teachers will depend on a number of issues, such as a) the accuracy of the answers sent to the tool; b) the complexity of the dialogue with the tool after the first answer; c) the level of applicability of the answer to the real context; and d) the final overall experience as user companion. Watson Analytics, in the context of this educational framework was useful to further analyze basic interaction data from a large dataset, over a 10-year period. The suggested draft framework provides a number of insightful answers to a generic set of questions, which is the most suitable to normalize the dialogue between the tool and the dataset for that along period that involves various teachers, students, platforms, educational models and administrative staff.

At present, the adoption of this framework relies on the data collection phase across multiple Learning Management Systems, so that the user information can be integrated into a single, normalized data set. Due to the structure of the educational system in China, the use of a single data base seems not possible, and the solution leans on the right combination of services to collect and harmonize the data from various sources and formats before the actual analysis. After this phase is achieved, the use of the framework and Watson Analytics will improve in time and depth.

From the technological perspective, future challenges are the development of an IMS LTI [27] based infrastructure that will allow the integration of the LMS and IBM Watson Analytics. Such development would enable question driven analytics to be launched right from the educational platform, and therefore would increase the usability of the framework.

Another promising future guideline is the research and development of live updates in the framework. With the current version, question driven analysis can be executed once the course has been finished, and it is therefore a posteriori analysis. Live questions open new scenarios where the teachers are enabled to adapt teaching strategies to the actual response of the students.

REFERENCES

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In Learning analytics (pp. 61-75). Springer New York.

Dawson, S., Bakharia, A., & Heathcote, E. (2010, May). SNAPP: Realising the affordances of real-time SNA within networked learning environments. Networked Learning.

Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. Journal of Learning Analytics, 4(2), 113-128.

Jovanovic, J., Gasevic, D., Brooks, C., Devedzic, V., Hatala, M., Eap, T., & Richards, G. (2008). LOCO-Analyst: semantic web technologies in learning content usage analysis. International journal of continuing engineering education and life long learning, 18(1), 54-76.

Marsh, J. A., & Farrell, C. C. (2015). How leaders can support teachers with data-driven decision making: A framework for understanding capacity building. Educational Management Administration & Leadership, 43(2), 269-289.

Wilson, A., Watson, C., Thompson, T. L., Drew, V., & Doyle, S. (2017). Learning analytics: challenges and limitations. Teaching in Higher Education, 1-17.

A Decision Support System for online teachers based on student similarities

Luis de-la-Fuente-Valentín*, Daniel Burgos
Universidad Internacional de La Rioja (UNIR)
{luis.delafuente, daniel.burgos}@unir.net
**Corresponding author*

ABSTRACT:

Online education is a rising market where the number of students is rapidly growing, forcing teachers to adopt tools and methodologies to support a larger number of students on a course. One of the most time consuming tasks is to compile the information needed for an effective tutoring process. Decision support systems automatically collect such information and present it to the teacher in the proper context, but are mostly focused on descriptive statistics. Assuming that students that behave in a similar way may receive similar feedback, the A4Learning tool measures similarity among students and visually represents it. This paper presents an empirical validation of A4Learning as a decision support system. In the presented case study the tool was deployed in two real courses with 48 students. The case study is analyzed in terms of usability, utility of the tool and accuracy of the tool in the classification task.

Keywords: online learning, decision support systems, similarity, data visualization

1 INTRODUCTION

Analytics tools provide data-driven support to teachers, fostering more effective teaching tasks. As presented by Dyckhoff et al. (2012), one possible approach is to collect students' activity and performance and show it to the teacher as an overview of 'what is happening on the course'. Data analysis is commonly used for predictive analysis, as in (Romero et al., 2013; Cambruzzi et al. 2015), which presents algorithms and techniques to increase accuracy while predicting students' final performance and dropout detection.

May (2011) suggests that Learning Analytics can be both descriptive and predictive. That is, instead of trying to foresee the future, descriptive analysis tends to explain the present by analysing the past. The relevant questions to answer are 'What happened?', 'Where was the problem?', and 'What actions are needed?' This paper presents a case study with A4Learning, a tool that supports tutoring tasks by providing descriptive information based on a visual analytics approach and on the similarity between students' behavioural patterns.

This paper is organized as follows. Section 2 presents the visualization tool subject of analysis. The methodological procedure used in the presented case study is detailed in Section 3, while Section 4 is devoted to present and discuss the results. Finally, Section 5 draws the conclusions of the study.

2 THE VISUALIZATION TOOL

A4Learning (de-la-Fuente-Valentín, Burgos & Crespo, 2014) is a decision support system that analyses student similarities and relates such analysis with the scores that students obtain at the end of the course. Two students are considered similar if they produce similar event logs in a given period. The similarity measurement is based on the type of event and the number of times each event type is repeated. For example: a student, Alice, is enrolled in a programming course. During the previous course edition (with similar duration, and the same learning activities and available tools), a monitoring system captured the students' activity, and is now capturing Alice's activities. Thus, Alice's activity pattern is being compared with that of students on previous courses and, for each historic student, says how similar Alice is to each of them.

The similarity information is visually related to the obtained score as follows: students from previous courses are grouped according their obtained score (0-1, 1-2, etc.). A4Learning calculates Alice's similarity with a group as the average similarity with all of the students in that group. The similarity value determines the colour (e.g. dark colours are for groups that are more similar). Figure 1 shows the resulting visualization, revealing that Alice behaves similarly to those students who scored from five to eight. A4Learning is not estimating Alice's score, but is showing the score of similar historic students.

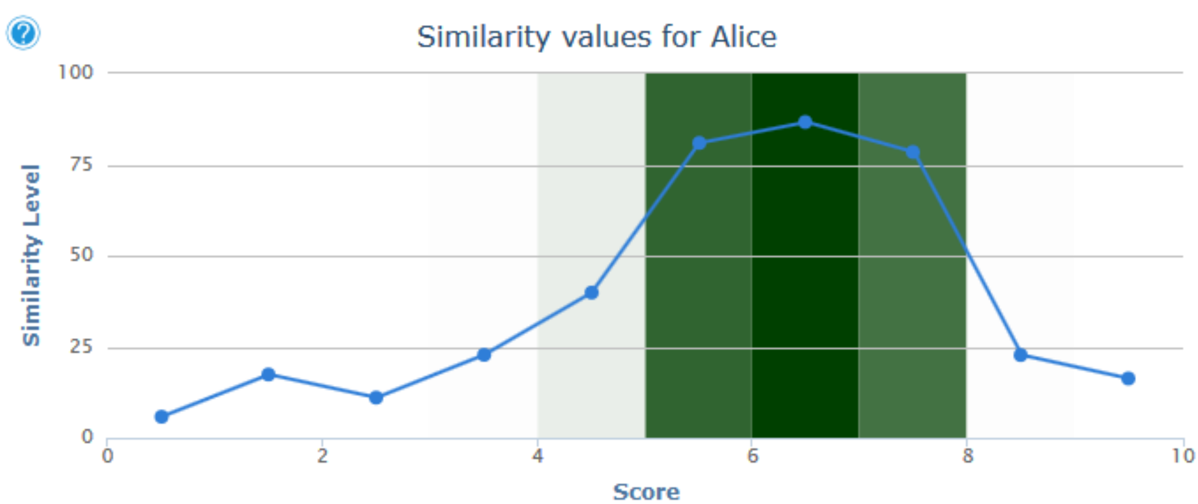


Figure 1: Student-centred view of similarities

More information about the A4Learning backend and development process can be found at (de-la-Fuente-Valentín et al. 2015, de-la-Fuente-Valentín & Burgos, 2014).

3 EXPERIMENTAL METHODS

3.1 Research questions

The case study was guided by five specific research questions. Firstly, consider A4Learning as a decision support system that provides a visual representation that classifies the students by students' expected results. The first two questions search the relationship among such classification and the actual results.

- **[RQ1] To what extent do the classifications match the actual results?**
- **[RQ2] To what extent do the classifications match the teachers's beliefs?**

The last three questions emphasizes that an important aspect to consider is the impact of the tool on the teachers' workflow. That is, if they have included or excluded tasks in their daily work:

- **[RQ3] What is the perceived usefulness of A4Learning?**
- **[RQ4] Does A4Learning use result in teacher actions that would not happen without the tool?**
- **[RQ5] Do the teachers understand and know how to use the visual information and the interface options?**

3.2 Experimental methodology

A4Learning was deployed in two online courses, Web projects management and Web services administration, both part of the same master's programme. These are fully online courses where during the four weeks of the course the students complete several activities and submit them by the last day of the course. The grades of these activities are weighted at 40% of the total grade. A final (face-to-face) examination provides the remaining 60% of the total grade. The 48 students took both courses. Both courses start and finish at the same time and were taught by the same teacher.

The courses lasted four-weeks, but the experimental setting was under observation during the last two weeks. This was planned this way to avoid the cold-start effect the tool was set up at the beginning of the course. Then, the teacher assisted to a training session where the researchers explained the characteristics and functionality of the tool.

3.3 Data capture methods

The researchers observed the case study during and after the course, collecting the following data:

- During the course

- Quantitative data: human-estimations (i.e. an estimation of the student's grade, given by the teacher with a simple slider interface and based on his/her beliefs); machine classifications (i.e. classifications made by the software with no human intervention).
- Qualitative data: email, personal communication with the teacher and reported comments in the bug tracking system

- After de course
  - Quantitative data: actual students' grade; platform logs of actual use.
  - Qualitative data: a questionnaire filled by the teacher with Likert-scaled, multiple choice, and open text questions. There were questions about usability, utility, precision and their general opinions of A4Learning. Also, the researchers had an interview with the teacher.

4 RESULTS AND DISCUSSION

4.1 Utility and usability

The analysis of the platform logs takes into account the number of times that the teacher logged in to A4Learning and which part of the tool she accessed. This usage analysis does not mean anything by itself, but helps with understanding the teacher's perceived utility. The analysis shows that the teacher used A4Learning daily, with more emphasis in the first day of the week. From the analysis of the logs and the questionnaire answered at the end of the course (summarized in Table 1), the following conclusions can be made to provide an answer to RQ3 and RQ4:

Table 1: Summary of utility questions

	Question	Answer
1	How often did you use the tool?	Most of the times I worked in the supported courses (answer from select box)
2	About the information given by A4Learning	I could reach the information by myself, but A4Learning makes the task more agile (answer from select box)
3	When you used the tool, what was your purpose?	Obtain information from the students (answer from select box)
4	Did you decide to actively support any student due to A4Learning information?	*Yes, some of the students* (answer from select box)
5	If your previous answer was 'yes', explain what type of support.	It was easy to find inactive students. I called them to understand what was happening.
6	Choose the reason for your support action.	I supported the student because A4Learning warned me about a situation I would have not found by myself. (answer from select box)

7	For what task did A4Learning support you?	To find students with low participation (answer from select box)
8	Did you integrate A4Learning into your daily workflow?	No, I did not.
9	Would you like to use A4Learning in future courses?	No, because in this case all the graded activities are delivered at the end of the course, and I do not know if the activity is enough to classify students. It would be preferable to use it in courses with continuous submissions.

- The teacher found useful information while using the tool, and used the collected information in her daily task. Derived from the A4Learning usage, she detected some situations and called the students involved (which is the ultimate role of the teacher).

- The teacher recognizes that she could obtain the same information from other sources, but A4Learning eased the task.

- Despite the number of page views from the statistical usage analysis (a minimum of 30 page views per day), the teacher did not consider the tool to be integrated into her workflow.

On the questionnaire, the teacher understood A4Learning as a tool that 'allows you to see the result that a student may have, taking into account students from previous courses that behaved similarly'. At the interview, the teacher also acknowledged 'I know that A4Learning also considers odd cases, because a student from previous courses may also have the same odd behaviour'. That is, she recognized that the estimation does not consider the amount of activity, but the similarity measures. Such quotations from the questionnaire and the interview reveal that the teacher actually understood the nature of the tool and provides an answer to RQ5.

4.2 Accuracy of the classification

The accuracy of the estimations (answer to RQ1) was measured by taking into account the classification made by the machine (with no human interpretation), the estimations by the teacher (with the support of A4Learning) and the actual obtained scores. The analysis was based on the success rate on classification.

The success rate was calculated for the 29 human estimations received, with success in eight of them (27% success rate). The machine estimation for these 29 cases succeeded in seven of them, and none of the cases did the teacher and the tool succeed at the same time. When comparing the success of estimations between human and automatic estimations (see Figure), it was noticed that the teacher had a better success ratio for students who passed the course, while the automatic estimations succeeded in those cases where the student did not take the final examination. In other words, A4Learning behaves as an early warning system for those students at risk, while for successful students A4Learning is a decision support system where human contextualization is required.

As an answer to RQ2, the teacher recognized that 'In general terms, the A4Learning estimations matched my opinion, built upon my conversations with the students'. The teacher's comments also emphasizes one of the main characteristics of A4Learning as a visual analytics tools: the need for the human interpretation of machine results to contextualize the data. Regarding this fact the teacher said, 'In some cases, I found severe risk students, but I know their personal circumstances and I know that they will do a good job'.

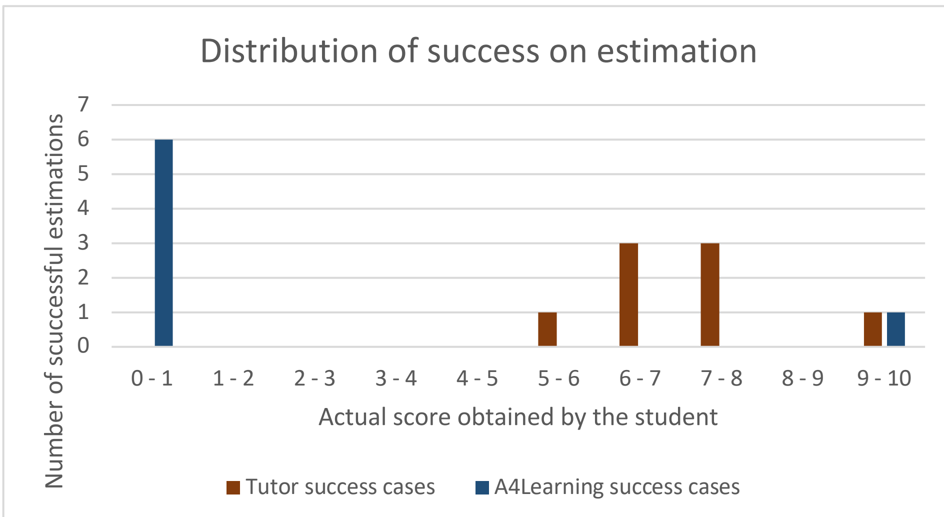


Figure 2 : Distribution of success on estimation

4.3 Methodological and adoption considerations

In the presented case study, the A4Learning tool is used during the last two weeks of a four-weeks course. The short duration of the case study was required due to the restrictions of the scenario, while it might affected the way the teachers perceived the tool. That is, two weeks might be enough to study how did the teacher understand the tool, but might not be sufficient to observe the tool usage in the long term.

Regarding issues that may affect the tool adoption, the researchers are concerned on the following aspects:

- Teachers should clearly understand the visual representation presented by the tool. In other words, the teacher should understand the visualization as something like "students behaving like Alice averaged 8 points", instead of "Alice obtains 8 points". A misunderstanding would hinder the adoption of the tool, since the tool is intended to support teachers and not to automate evaluation tasks.

- Teachers might be biased by the visualization and (consciously or not) change their opinion on a certain student and give him/her a biased feedback, or a biased evaluation.

- As pointed out by the teacher on the last question of Table 1, the effectiveness of the tool might be affected by the pedagogical methods used in the course being supported. Students activity is modulated by graded activities and deadlines, and different configurations of deadlines may result in different activity patterns and affect the utility of the tool.

- Related to the previous one, the tool assumes that previous editions of the course were taught with the same pedagogical methods and similar chronological planning. Courses with many changes in different editions might not be compatible with A4Learning.

All these aspects may affect the effective adoption of A4Learning and therefore should be subject of future research.

5 CONCLUSSIONS

This article presents a case study of A4Learning, deployed in an authentic learning scenario where one teacher provides support to 48 students. The tool, aimed at helping the teacher in tutoring tasks, makes use of similarity metrics to compare students with those from previous courses, observing how they performed. The goal of the case study is the validation of the tool in terms of usability, utility and accuracy:

The analysis of teacher's usage of the tool, responses to the survey and responses to the interview show that he was able to identify cases that she would not have identified otherwise. Therefore, A4Learning was perceived as a useful tool that supports the tutoring tasks. The information presented by A4Learning is contextualized by the teacher, who was able to estimate the student's score. The teacher better estimated in those cases where the student finally passed the course. A4Learning behaved as an early warning system for those students at dropout risk. The data show that the teacher was able to provide better estimations for those students who finally passed the course, while A4Learning behaved as an early warning system for those students at risk, showing its potential as a supportive tool for teachers.

One important lesson learned is the need to separate description from estimation; that is, A4Learning is a descriptive system, and this fact must be understood by the end user in order to interpret its results. Another lesson learned is that the pedagogical methods of the course have a great influence on the effectiveness of the tool and the teachers' intention to adopt the tool as part of his workflow. Furthermore, the analytical approach of the tool requires few changes in the chronological planning of the course from one edition to another, which is not always the case. This may hinder the actual adoption of the tool.

REFERENCES

Cambruzzi, W. L., Rigo, S. J., & Barbosa, J. L. (2015). Dropout Prediction and Reduction in Distance Education Courses with the Learning Analytics Multitrail Approach. J. UCS, 21(1), 23-47.

de-la-Fuente-Valentín, L., Burgos, D., & Crespo, R. G. (2014, July). A4Learning--A Case Study to Improve the User Performance: Alumni Alike Activity Analytics to Self-Assess Personal Progress. In Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference on (pp. 360-362). IEEE.

de-la-Fuente-Valentín, L., Pardo, A., Hernández, F. L., & Burgos, D. (2015). A Visual Analytics Method for Score Estimation in Learning Courses. J. UCS, 21(1), 134-155.

de-la-Fuente-Valentín, L, & Burgos, D. (2015). Am I doing well? A4Learning as a self-awareness tool to integrate in Learning Management Systems. Campus Virtuales, 3(1), 32-40.

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. Journal of Educational Technology & Society, 15(3), 58.

May, T. A. (2011). Analytics, University 3.0, and the future of information technology. EDUCAUSE Review Online.

Romero, C., López, M. I., Luna, J. M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. Computers & Education, 68, 458-472.

Measuring Impact of S3 Adoption on At-Risk Students

Shady Shehata
D2L Corporation
shady.shehata@d2l.com

ABSTRACT: Students' persisting to completion of educational goals is a key gauge of student success, and therefore institutional success. Advancements in data analysis and predictive modeling have tremendous potential to improve student success by enabling colleges and universities to build powerful predictive models that predict student behavior.

In this paper, we first describe Student Success System (S3) that can measure student performance starting from the first weeks of the semester. S3 is an Early Intervention System that empowers institutions with predictive analytics to improve student success, retention, completion, and graduation rates. S3 provides educators with early indicators and predictions of student success and risk levels. S3 provides interactive visualizations to highlight patterns and indicators about the student and their position relative to the course expectations and to other students within the class. Secondly, we measure the instructor adoption of the student page visualizations that enable instructors to track student performance to make interventions. Thirdly, we analyze the correlation between the adoption of student visualizations and percentage of at risk students. Finally, and most importantly, we describe the adoption process of S3 by the institution including the role of stakeholders. We present the findings and the lessons learned of the adoption process.

Keywords: Learning Analytics, Data Mining, Machine Learning, Predictive Modeling, Predictive Analytics, Visualizations, Adoption, Student Performance, Correlation, Student Success.

1 INTRODUCTION

Predictive modeling allows faculty and administrators to understand why some students persist and others do not. This knowledge can be leveraged into positive action through targeted intervention programs. In education, predictive models for identifying at-risk students were first introduced by (Campbell, 2007). Similar work has been underway at a variety of institutions, including Marist College, Capella University, University of Phoenix and Rio Salado College (Lauría et al., 2013; Wolff et al. 2014; Hlosta et al., 2014; Gilfus Education Group, 2012).

In this paper, we describe Student Success System (S3) that can measure student performance starting from the first weeks of the semester. S3 is developed by D2L Corporation and is in production. S3 provides educators with early indicators and predictions of student success and risk levels. Predictions generated by S3 are based on predictive models that are created by applying machine learning algorithms on historic course data (usually prior offerings of the same course for which predictions are to be generated). The predictive analysis is adaptable and customizable to the instructional approach of each course, as well engagement and achievement expectations. The system provides weekly predictions of student success levels within their courses in the form of a success index. The success index is designed to let the instructors visualize and compare key factors and to design interventions. Section 2 describes the Student Success System overview and focuses on student page visualizations that track student performance. Student page visualizations enable instructors with the analytics to help instructors to design their interventions. Section 3 describes data analysis of adopting S3 visualizations with measuring correlation between the access count of student page visualizations and percentage of at-risk students in each course. Section 4 presents the process of S3 adoption by the institution with demonstrating the roles of the stakeholders and the lesson learned in the adoption process.

2 STUDENT SUCCESS SYSTEM

S3 provides interactive visualizations for instructor(s)/advisor(s)/educator(s) to highlight patterns and indicators about the student and their position relative to the course expectations and to other students within the class. There are no visualizations available to students yet. These visualizations allow instructors access to new and powerful information never previously at their disposal. These visualizations help instructors with analytics that detect student success and identify the missing activities that are required for the students to make to improve performance.

2.1 Overview

Instructors can monitor the status of all students in terms of their predicted success; At Risk, Potential Risk, and Successful. The levels are determined based on thresholds on the predicted grade. The defaults are: 0%-60% for At-Risk, 60%-80% for Potential Risk, and 80%-100% for successful. These instructor-specified thresholds can be configured for each course during the setup of the predictive model. The value of the success index is determined based to the predicted grade.

2.2 Student Page Visualizations

As shown in Figure 1, instructor can choose to drill down on an individual student to gain the insights the instructor needs about the individual student, so that the instructor can design a personalized intervention. The student page has the most important visualizations that track student engagement.

2.2.1 Student Profile

On the top-left side in Figure 1, basic profile information about the student is presented. This includes the student picture, name, and campus Id. If the institution permits access to addition data elements from the Student Information System (SIS), they will be presented.

2.2.2 Course Timeline

On the top–right side in Figure 1, a timeline chart shows the weekly success index, indicated by its value on the y-axis and color. It provides at-a-glance view of the student trend-line.

2.2.3 Win-Loss Chart

On the bottom-left in Figure 1, a win-loss chart shows how the success index is designed as a combination of predictors. The chart shows the student success based on each of these predictors and how they compare to the mid-range expectations. The vertical reference line in the middle of the chart corresponds to the middle point within in range of Potential Success (if the range is 60%-80%, the middle point is 70%, which corresponds to the value 7 for the success index). The success indicator for the student along each of these indicators is compared to this reference point, leading to either a “win” where the bar falls on the right side of the line or a “loss” where the bar is on the left side of the line. The “loss” indicators point to expectation gaps where improvement can be made.

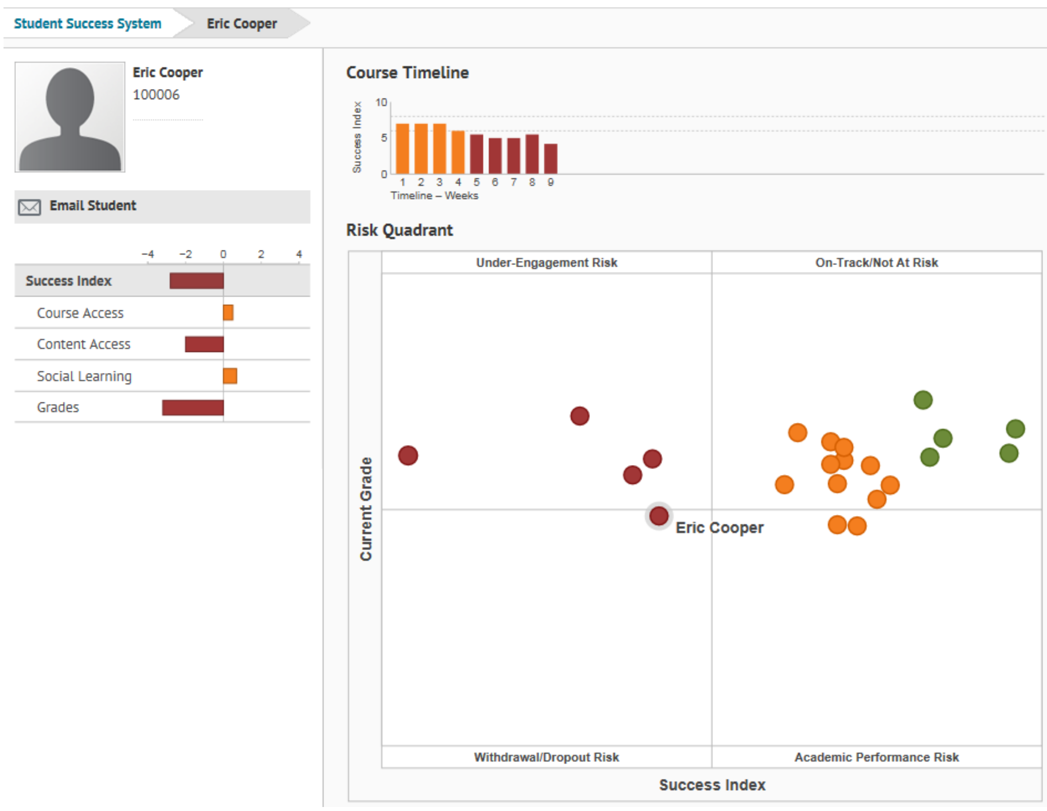


Figure 1: Student page visualizations

2.2.4 Risk Quadrant

The Risk Quadrant positions the student within the class based on the success index and the current grade as shown in Figure 1. The success index provides the overall predicted outcome that is primarily based on engagement. The current grade provides an overall measure of performance based on the calculated grade to-date. Each dot on the chart is a student. The current student is highlighted.

In the next two sections, we will discuss Student Success adoption in two levels. The first level presents data analysis of S3 technology adoption with measuring the correlation between S3 visualizations adoption and percentage of at-risk students. The second level and most importantly focuses on the adoption process by the institution. The adoption process includes the interactions among the stakeholders and their roles in the process. We focus on presenting the findings and the lessons learned of the adoption process by the institution.

3 DATA ANALYSIS OF S3 TECHNOLOGY ADOPTION

We calculate the usage of all S3 visualizations across institutions in the time frame of January to December of 2017. The usage is presented by "Access Count" which is number of visits to all S3 visualizations as in Figure 2.

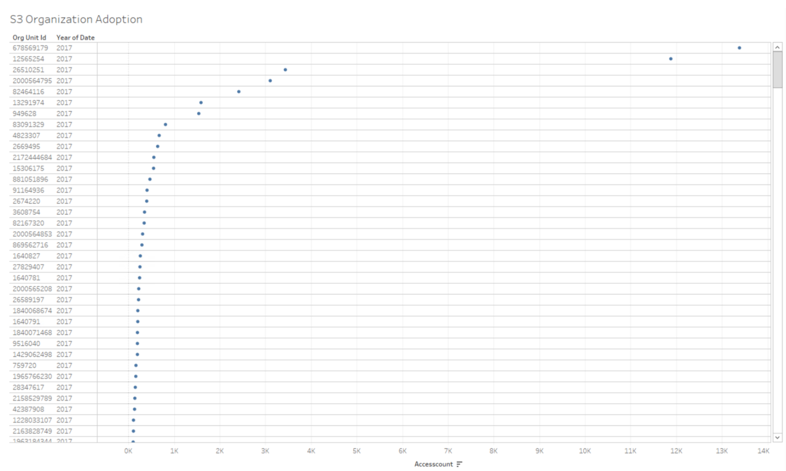


Figure 2: S3 Adoption across institutions

In Figure 2, the x-axis is the access counts of all the visualizations in S3 including visualizations for building predictive models, student dashboard, student page visualizations, social learning, and grades visualizations. The y-axis is the organization id. As shown in Figure 2, the access count reaches the value of 14K visits within institutions that are adopting S3.

3.1 Student Page Visualizations Adoption

In this subsection, we focus on calculating the adoption of the student page visualizations of Figure 1 since it provides a drill down on an individual student performance. The visualizations in the student page provide the insights the instructor needs about each individual student, so the instructor can design a personalized intervention. In Figure 3, we measure the access count of student page visualizations by instructors on ten courses that are in the top S3 usage offered in the same time frame of January to December of 2017. The access count is the number of visits to the student page.

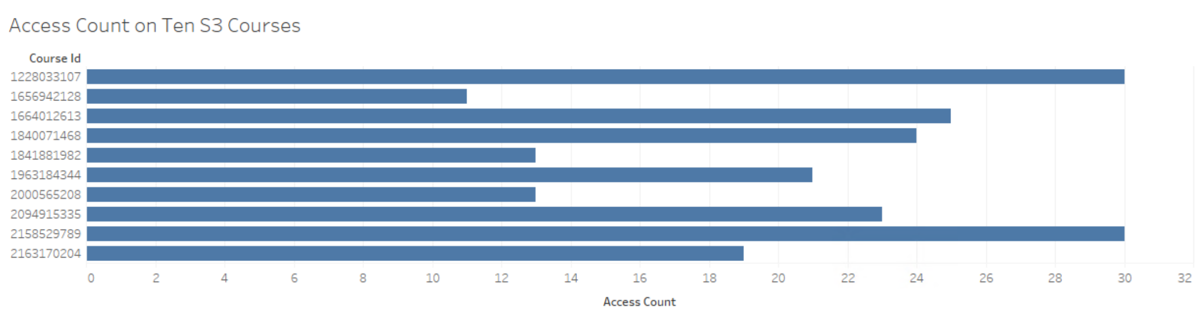


Figure 3: S3 Access count for student page visualizations on ten courses

3.2 Correlation Analysis Results

We calculate the percentage of students that are at risk in courses offered in the time frame of January to December of 2017 based on the final grades of students. Student is at risk when he/she has a final grade value that is less or equal than 60. For ten courses that are in the top of S3 usage, we calculate the correlation between access count of the student page visualizations and percentage of at risk students.

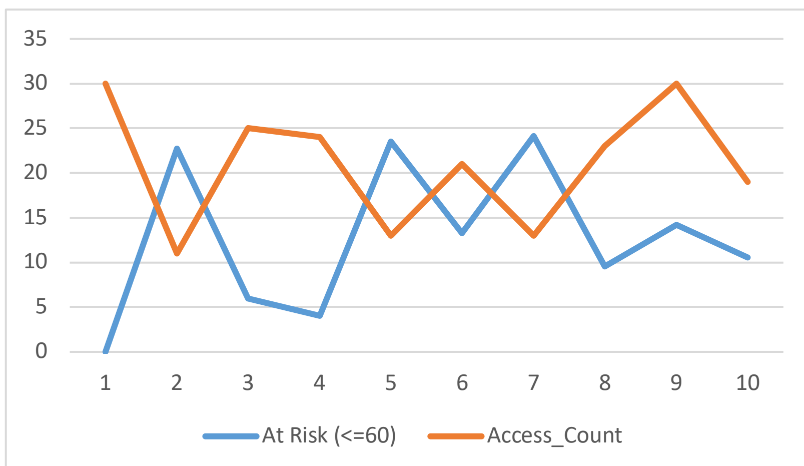


Figure 4: Relationship between student page visualization and percentage of at risk students

As shown in Figure 4, for the 10 courses, we observe there is a strong negative correlation by the value of -0.829110948 between student page visualizations access counts and percentages of students that are at risk. It's observed that the more visits by instructors to student page visualizations, the less value of student percentage that are at risk. Student page visualizations track student engagement and performance and demonstrate the educational gaps (i.e. student needs to read more content to be successful) that students need to make to improve their performance. One interpretation is that visiting these visualizations frequently by instructors may present the intention by instructors to look for insights and analytics to help instructors in designing their interventions with their students and the more interventions occurring, the less number of at-risk students.

4 S3 ADOPTION PROCESS BY INSTITUTION

In this section, we focus on demonstrating the process of adopting S3 by the institutions. The adoption process presented in this section includes different personas and their roles of each persona in the adoption process and the lessons learned.

4.1 S3 Pilot Project

Some institutions like to begin piloting Student Success System on specific courses by selecting pilot faculty based on evidence of iterative student intervention where Instructor interest was gauged. In this case, the institution dedicates significant resources to evaluate how models were: configured and selected, behaved over time (semester), and impacted individual students. For each pilot course, model criteria and historical course data was gathered, domain configurations were created, and models were built in simulation mode. A model matrix was then created for each pilot course as shown in Figure 5. Each model matrix laid out between 18 and 25 possible model configurations along with aggregate error measures (Mean Squared Error and Average Percent Correct). A final metric attached to each simulated model was the percentage of students ultimately receiving an unsuccessful outcome (grade of D or F) that had an individual error greater than 10%. This final metric was important to help identify potential models that may have adverse individual impact. After the matrices were created, each matrix was presented to the pilot faculty member, and up to three models were selected for more detailed review.

Model	Course	Historical Courses	Bands	Course	Content	Social	Grades	Prep	APC	MSE	>10% RED
One	ENGL 101: Freshman Composition S13	ENGL 101: Freshman Composition F12 ENGL 101: Freshman Composition S12	0-60% 60-75% 75-100%	YES	YES	NO	YES	YES	97.8%	8.9%	100%
Two	ENGL 101: Freshman Composition S13	ENGL 101: Freshman Composition F12 ENGL 101: Freshman Composition S12	0-60% 60-75% 75-100%	YES	NO	NO	YES	YES	57.6%	18.2%	25%
Three	ENGL 101: Freshman Composition S13	ENGL 101: Freshman Composition F12 ENGL 101: Freshman Composition S12	0-60% 60-75% 75-100%	YES	YES	YES	YES	NO	83.8%	6.1%	20%

Figure 5: Sample S3 model matrix

Aggregate error measure and other associated metrics can be very helpful in delineating between potential models. However, some institutions found that aggregate measures alone were not sufficient to select a model for high-risk decision making for them. Some institutions piloted S3 it is important to look at how different model configuration affected individual students. Therefore, after faculty selected up to three models for more detailed review, and student level previews were generated as shown in Figure 6. These renderings allow for a student-level view of how each model acts across the weeks in a semester, providing an additional level of information to help the instructor and S3 administrator make the best model selection. Of key importance at this juncture is to evaluate various model configurations for base-level disparate student impact. Additionally, the individual student-level preview allows for an easy visual cue for many instructors about when a model stabilizes to help instructors design their intervention strategy. After a final model selection was made, the S3 Administrator would rebuild the model in “live” mode, rather than simulation mode, and confirm that the model was set in the S3. S3 would generate new predictions automatically every seven days.

	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	Final	MSE (grade)	MSE (SI)	APC
Student A	8	6	6	8	8	8	7	7	8	8	8	7	7	7	7	7	6.7%	6.5%	81%
Student B	6	7	7	7	6	6	7	7	7	6	6	7	6	6	6	6	6.7%	6.9%	57%
Student C	7	8	8	8	7	7	8	7	8	7	7	8	8	7	7	7	8.2%	6.2%	38%
Student D	7	8	9	8	8	9	8	8	8	8	8	8	9	8	8	6	3.7%	4.9%	14%
Student E	6	7	7	7	7	7	7	7	7	7	6	7	6	6	6	7	5.3%	3.6%	62%
Student F	7	6	5	5	5	4	4	4	4	5	4	4	5	6	6	6	9.1%	12.5%	76%
Student G	7	8	9	7	8	8	8	8	8	8	8	8	8	7	7	7	5.9%	8.2%	96%
Student H	6	5	5	6	6	6	7	6	7	6	7	7	7	6	5	5	13.8%	12.1%	19%
Student I	7	8	8	7	7	7	8	8	8	7	7	8	8	7	7	8	5.6%	6.9%	38%
Student J	7	8	9	8	8	8	8	8	8	8	8	8	9	8	8	9	5.9%	10.9%	86%
Student K	6	5	5	5	5	5	4	4	4	4	3	3	2	2	2	0	38.4%	41.2%	67%
Student L	6	7	6	6	6	7	7	7	7	7	7	7	7	7	6	7	8.0%	7.9%	81%
Student M	7	8	8	8	7	7	7	7	8	8	8	8	8	7	7	7	3.5%	6.5%	61%
																Min	3.5%	4.9%	14%
																Max	38.4%	41.2%	96%
																Avg	9.3%	10.6%	63%
																SD	9.1%	9.6%	22%

Figure 6. Sample S3 student-level preview

4.2 Lessons Learned

One lesson learned is that the communication between S3 Administrator and instructors are vital for the adoption process to be successful. Choosing among the many generated predictive models will require the knowledge of both S3 Administrator and Instructor. After the instructor evaluated the model matrix for his/her course, followed by the student level-preview, the instructor, is advised to be in concert with the S3 Administrator, selected the model that he/she felt would best allow his/her to support the students in his/her class. Instructors relied on the error measures and other metrics, along with which model would supplement their intervention and support strategies. The ultimate model selection was always left to the discretion of the course instructor since the instructor knows the best on which tools have been used in his/her course.

Another lesson learned is that continuous evaluation and maintenance strategy for the running predictive models should be planned from the start to improve the predictive models' quality overtime. Continuous evaluation of predictive models helps the instructor to understand the reasoning behind the generated predictions to build the trust between the stakeholders and S3 predictions. The lack of predictive models' evaluation can lead to losing the trust of the learning analytics solution which will have negative impact on adopting learning analytics solution. The output of the evaluation process should be incorporated to S3 system to ensure that the predictive models performance is improved over time.

Other lessons learned is that creating a group of selected instructors who are interested in running S3 on their courses is important for the adoption process in the institution. Overtime this group educates more stakeholders by providing their experiences about dealing with S3 to other members. The group expands to add more individuals to run S3 on more courses and adopt S3 by more instructors. We found that this approach helps to scale the adoption of S3 in the institutions overtime.

REFERENCES

Campbell, J. P. (2007). Utilizing student data within the course management system to determine undergraduate student academic success: *An exploratory study*. ProQuest.

Gilfus Education Group. (2012). Academic analytics, New eLearning diagnostics. Retrieved from: http://www.gilfuseducationgroup.com/academic-analytics-new-elearningdiagnostics

Hlosta, M., Herrmannova, D., Vachova, L., Kuzilek, J., Zdrahal, Z. and Wolff, A. (2014). Modelling student online behaviour in a virtual learning environment, *Workshop: Machine Learning and Learning Analytics at Learning Analytics and Knowledge (LAK).* Indianapolis.

Wolff, A., Zdrahal, Z., Herrmannova, D., Kuzilek, J. and Hlosta, M. (2014). Developing predictive models for early detection of at-risk students on distance learning modules. *Workshop: Machine Learning and Learning Analytics at Learning Analytics and Knowledge (LAK)*. Indianapolis.

Personalising feedback at scale: approaches and practicalities

Lorenzo Vigentini
UNSW Sydney
l.vigentini@unsw.edu.au

Danny Y.T. Liu
University of Sydney
danny.liu@sydney.edu.au

Lisa Lim
The University South Australia
Lisa.Lim@unisa.edu.au

Roberto Maldonado
University of Technology Sydney
Roberto.Martinez-Maldonado@uts.edu.auEmail

ABSTRACT: In the context of moving towards user-centred analytics, one of the most impactful practices for improving learning and teaching is the provision of effective and timely feedback of, and for, learning. This workshop aims to bring together scholars and practitioners to find a common ground for showcasing interesting examples of effective feedback, demonstrate the use of supporting tools, and explore what and how data can be used to improve the process and richness of feedback for both learners and educators. Key outcomes will be a better understanding of tools, approaches and existing cases of good practice which will foster discussion and collaboration in the LA community.

Keywords: personalization, feedback, user-centred analytics

1 INTRODUCTION

1.1 Background

User-centred analytics benefits greatly from impactful practices such as the provision of effective and timely feedback of and for learning (Brown & Knight, 1994; Hattie & Timperley, 2007; Hounsell, 2003; Sadler, 1989). Although feedback is essential in promoting autonomy and self-regulation in higher education (Black, Harrison, & Lee, 2003; Nicol, 2010; Sadler, 2010), it is often the lowest rated aspect in terms of satisfaction from graduate satisfaction surveys like the NSS, CEQ, SES, etc. (Krause, Hartley, James, & McInnis, 2005; McDowell, Smailes, Sambell, Sambell, & Wakelin, 2008; Radloff, Coates, James, & Krause, 2011; Rowe & Wood, 2009; Williams & Kane, 2008). More importantly, while assessment practices have received considerable attention over the past decade: examples as REAP, SAFE, 'Transforming Assessment' projects as examples - (Carless, Salter, Yang, & Lam, 2011; Crisp, 2011; Nicol, 2009), the focus on feedback to students has remained relatively scarce (Higgins, Hartley, & Skelton, 2002; Orsmond, Maw, Park, Gomez, & Crook, 2013; Rowe & Wood, 2009). At the same time, the growing interest in learning analytics (LA) has brought to the forefront the potential of using behavioural, engagement and other sources of data captured from learning and teaching activities to be able to improve the timeliness, relevance, and effectiveness of feedback. The *personalisation* of feedback (using LA) has become a sort of holy grail for educators aspiring to improve their students' learning and satisfaction (Bienkowski, Feng, & Means, 2012; King, Kinash, Kordyban, & Pamenter, 2014).

Students and educators do not hold the same perception of what constitutes quality feedback (Carless, 2006; Forsythe & Johnson, 2017; Hounsell, McCune, Hounsell, & Litjens, 2008; Lizzio & Wilson, 2008; Pitt & Norton, 2017). In most cases, the idea of providing feedback is reduced to a summative, corrective and transmissive process, which gives a final judgement on students' submitted assignments (Nicol, 2010; Weaver, 2006). In order to improve the process, some researchers (Forsythe & Johnson, 2017; Pitt & Norton, 2017) have started to reconsider the impact of (or lack of) feedback as currently implemented in HE and, instead, focus more on the constructive value of a dialogic approach in which both giving and receiving feedback are considered more holistically (Forsythe & Johnson, 2017; Nicol, 2010; Pitt & Norton, 2017; Poulos & Mahony, 2008). In today's global higher education climate of massification and diversification, another important consideration involves enabling teachers to provide relevant feedback effectively and efficiently. Although LA have made tangible connections with critical aspects that can strongly shape learning, such as learning design and self-regulation, the provision of feedback to students has been relatively neglected (Liu, Bartimote-Aufflick, Pardo, & Bridgeman, 2017; Pardo, 2017). This is despite the affordances of LA to leverage the generation of theoretical and technical mechanisms for understanding and improving learning by "informing and empowering instructors and learners" (Siemens & Baker, 2012). To allow this to happen, teachers need concrete tools and approaches to bridge the gap between LA research and classroom practice. LA systems are starting to support teachers with means to provide rich feedback beyond typical early warning messages (e.g. SRES, Ontask, (Tempelaar, Rienties, & Giesbers, 2015), but it is clear that there is a need and appetite in the LA community of research and practice to further explore data-informed student-centred pedagogies to provide feedback at scale.

1.2. Scope of the workshop

This workshop brings together scholars and practitioners to explore interesting examples of effective feedback, demonstrate the use of supporting tools, and explore what and how data can be used to improve the process and richness of feedback for both learners and educators. The workshop has three primary goals:

- Provide a multidisciplinary theoretical foundation for practitioners and researchers in LA for the effective provision of data-informed feedback practices in HE;
- Showcase current implementations of tools and methods which enhance feedback practices, especially around personalisation;
- Promote reflection on both pedagogical and technological approaches to improve feedback practices targeted at the improvement of student learning and their ability to self-regulate learning.

2. ORGANISATION DETAILS

This half-day workshop invited contributions on topics such as tool(s)/approach(es) to personalise feedback, the implementation process (e.g. infrastructural, staff capacity, etc.) and challenges and successes (as well as failures). After receiving several submissions, these were peer reviewed leading to six accepted papers representing work conducted in Australasia, Europe and the USA. Four different tools/systems are presented including both open source and commercial and showcasing case studies from several universities. Here is the list of papers included:

- *Liu et al.* Trojan horse analytics: Hooking educators on personalized feedback at scale at multiple Australian universities
- *Lim et al.* Combining technology and human intelligence to provide feedback and learning support using *OnTask*
- *Bucic et al.* Juggling system and data streams to personalize feedback: the case of a large first year course
- *Moxley & Bennington* Actionable Analytics at MyReviewers for Administrators, Instructors, Students, and Researchers
- *Loftus & Madden* Probabilistic Graphical Models as Personalised Feedback
- *Essa & Aghababyan* Personalizing Non-Cognitive Feedback for Learning and Skill Formation

2.1. Who is this workshop for?

Those who wish to understand and apply principles of feedback of and for learning. Given the explicit multidisciplinary nature of the workshop we expect that it will provide an opportunity to discuss and share innovations, impact on learning, and explore future directions in the application of learning analytics (LA) to personalisation of feedback. Likely interested participants:

- Educators/teachers and researchers
- Technologists and educational developers
- Learning scientists and data scientists/analysts
- Academic managers
- and anyone else interested in personalisation of learning and teaching

2.2. Proposed workshop activities

After a brief introduction and conceptualisation of the workshop, in the first half of the workshop a rapid series of short presentations will provide a backdrop and provocation to think about ways in which we normally provide feedback in HE and present samples of tools and cases in which an approach has been well received or successful (or discuss reasons for failure).

In the second half, breakout groups will be guided with a semi-structured approach to discuss key themes and issues surfaced during presentations (e.g. use of data-informed feedback by students, types of feedback made possible through data, challenges of faculty professional learning, data sources needed for personalisation, etc.).

A website will be created to provide access to all contributions and presentations as well as a summary from the organisers after the workshop. The workshop will provide an avenue to continue the conversations beyond the session and open opportunities for further collaborations.

3. INTENDED OUTCOMES FOR PARTICIPANTS

We expect a range of presentations that will cover practical, evidence-based approaches to personalising data-driven feedback at scale. Participants will be able to:

- Obtain a broad perspective of different approaches to using data for personalising feedback
- Enhance their understanding of the forms of feedback that could improve student learning
- Gain an appreciation of the range of contexts where feedback can be valuable, and how data can inform these
- Discuss cases, issues, and potential solutions to implementing LA-enhanced feedback practices
- Connect with researchers and practitioners working to provide personalised feedback, yielding opportunities for collaborating on approaches and tools across attending institutions.

After the workshop, given the commitment to further collaborations, contributors will be invited to consider more substantial submissions with the intention to collate the works into either a special issue of journal, or CEUR proceeding or an edited book on the topic.

4. REFERENCES

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology*, *1*, 1–57.

Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice*. McGraw-Hill Education (UK).

Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. Psychology Press.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, *31*(2), 219–233. https://doi.org/10.1080/03075070600572132

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, *36*(4), 395–407. https://doi.org/10.1080/03075071003642449

Crisp, G. (2011). *Rethinking assessment in the participatory digital world – Assessment 2.0* (National Teaching Fellowship Fil Report).

Forsythe, A., & Johnson, S. (2017). Thanks, but no-thanks for the feedback. *Assessment & Evaluation in Higher Education*, *42*(6), 850–859. https://doi.org/10.1080/02602938.2016.1202190

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Higgins, R., Hartley, P., & Skelton, A. (2002). The Conscientious Consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, *27*(1), 53–64. https://doi.org/10.1080/03075070120099368

Hounsell, D. (2003). Student feedback, learning and development. *Higher Education and the Lifecourse*, 67–78.

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development*, *27*(1), 55–67. https://doi.org/10.1080/07294360701658765

King, C., Kinash, S., Kordyban, R., & Pamenter, J. (2014). Personalising student learning through education. Bond University. Retrieved from http://epublications.bond.edu.au/tls/87

Krause, K. L., Hartley, R., James, R., & McInnis, C. (2005). *The first year experience in Australian universities: Findings from a decade of national studies*. Centre for the Study of Higher Education, University of Melbourne Melbourne. Retrieved from http://www.cshe.unimelb.edu.au

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-Driven Personalization of Student Learning Support in Higher Education. In *Learning Analytics: Fundaments, Applications, and Trends* (pp. 143–169). Springer, Cham. https://doi.org/10.1007/978-3-319-52977-6_5

Lizzio, A., & Wilson, K. (2008). Feedback on assessment: students' perceptions of quality and effectiveness. *Assessment & Evaluation in Higher Education*, *33*(3), 263–275. https://doi.org/10.1080/02602930701292548

McDowell, L., Smailes, J., Sambell, K., Sambell, A., & Wakelin, D. (2008). Evaluating assessment strategies through collaborative evidence-based practice: can one tool fit all? *Innovations in Education and Teaching International*, *45*(2), 143–153. https://doi.org/10.1080/14703290801950310

Nicol, D. (2009). Assessment for learner self-regulation: enhancing achievement in the first year using learning technologies. *Assessment & Evaluation in Higher Education*, *34*(3), 335–352. https://doi.org/10.1080/02602930802255139

Nicol, D. (2010). From monologue to dialogue: improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, *35*(5), 501–517. https://doi.org/10.1080/02602931003786559

Orsmond, P., Maw, S. J., Park, J. R., Gomez, S., & Crook, A. C. (2013). Moving feedback forward: theory to practice. *Assessment & Evaluation in Higher Education*, *38*(2), 240–252. https://doi.org/10.1080/02602938.2011.625472

Pardo, A. (2017). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education*, *0*(0), 1–11. https://doi.org/10.1080/02602938.2017.1356905

Pitt, E., & Norton, L. (2017). 'Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assessment & Evaluation in Higher Education*, *42*(4), 499–516. https://doi.org/10.1080/02602938.2016.1142500

Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: the students' perspective. *Assessment & Evaluation in Higher Education*, *33*(2), 143–154. https://doi.org/10.1080/02602930601127869

Radloff, A., Coates, H., James, R., & Krause, K.-L. (2011). Report on the Development of the University Experience Survey. *Higher Education Research*. Retrieved from http://works.bepress.com/hamish_coates/79

Rowe, A. D., & Wood, L. N. (2009). Student Perceptions and Preferences for Feedback. *Asian Social Science*, *4*(3), 78. https://doi.org/10.5539/ass.v4n3p78

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*(5), 535–550. https://doi.org/10.1080/02602930903541015

Siemens, G., & Baker, R. S. J. d. (2012). Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (pp. 252–254). New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330661

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, *47*(Supplement C), 157–167. https://doi.org/10.1016/j.chb.2014.05.038

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, *31*(3), 379–394. https://doi.org/10.1080/02602930500353061

Williams, J., & Kane, D. (2008). Exploring the National Student Survey Assessment and feedback issues. *The Higher Education Academy, Centre for Research into Quality*.

Trojan horse analytics: Hooking educators on personalized feedback at scale at multiple Australian universities

Danny Liu
The University of Sydney
danny.liu@sydney.edu.au

Lorenzo Vigentini
UNSW Australia
l.vigentini@unsw.edu.au

Elsuida Kondo
The University of Melbourne
elsuida.kondo@unimelb.edu.au

Kevin Samnick
The University of Sydney
kevin.samnick@sydney.edu.au

Ruth Weeks
The University of Sydney
ruth.weeks@sydney.edu.au

ABSTRACT: University educators across the world are increasingly pressed for time and resources in the face of growing student numbers and needs. Learning analytics originally provided a promising solution which leveraged computing speed and data insight in order to better support students pedagogically and pastorally. However, the reality in many institutions is far from this. We contend that this may be due to a lack of educator-driven tools that address their (and their students') actual needs. Instead of analyzing and visualizing pre-existing warehoused 'big data', our experience in engaging educators with learning analytics suggests that personalized feedback that is driven by relatively 'small data' are key to strong adoption. Moreover, learning analytics tools that afford educators the ability to streamline and act on the entire data workflow (from capture to manipulation to action) are critical. A unique open source tool, the Student Relationship Engagement System (www.sres.io), developed at an Australian university, affords this. In this paper and the accompanying workshop, we will explore (i) the practical needs and contexts of educators and students that can be enhanced using data; (ii) the impactful operationalization of educator-driven learning analytics across several Australian universities; and (iii) collaborative developmental efforts between institutions.

Keywords: feedback; student support; personalization; faculty adoption; implementation; learning analytics; cross-institutional collaboration.

1 BACKGROUND

Reports from around the world suggest laggard adoption and implementation of LA by institutions and educators. Recent reports on Australasian LA adoption and implementation have highlighted that, as the primary implementers of any LA tool, educators need to be involved in designing LA approaches that "are sensitive to their environments, meeting and extending their pedagogical requirements, and ensuring flexibility" (Colvin et al., 2016, p. 19). In this context, a key need seems to revolve around actions that involve personal contact with their students, which balances the automation of computers with the personal approach of teaching (West et al., 2015). Notably, this report highlighted that educators "still have to make sure that it [communication and feedback] is personalised and meaningful for students", and that educators need LA tools with "some ability to modify it to their own requirements because each course and each cohort of students may differ" (p. 20).

2 SUPPORTING THE LEARNING AND TEACHING CONTEXT

The learning and teaching landscape in any institution, faculty, and indeed course or subject, is unique and influences uptake of any innovation, especially learning analytics (Ferguson et al., 2014). Several factors can impact adoption, but some are particularly pertinent to user-centered (i.e. educator- and student-centered) learning analytics:

- Faculty resistance to change and workload issues are examples of social and cultural context that need to be understood and addressed (Macfadyen & Dawson, 2012). This includes concerns around needing to adapt to new tools and approaches, and change existing practices.
- Depending on context, a large proportion of learning and teaching activities occurs outside the online space (and not just outside the confines of a learning management system), and often involves human interaction (West et al., 2015). This poses challenges for capturing and using data to support engagement and feedback.
- The lack of availability of tools that properly address the needs of educators and students (Colvin et al., 2016). This presents issues in moving educators along the learning analytics adoption pipeline, causing stalling or retraction of interest.
- A skills shortage to be able to analyze and interpret data, and a lack of professional development opportunities to address this in existing faculty (Gunn, McDonald, Donald, Blumenstein, & Milne, 2016). This assumes that the data, or at least the analyses, are more complex than faculty are used to working with.
- A lack of groundswell support and sharing that is driven by learning analytics users (educators and students) who have personally experienced tangible benefits.

From these challenges, it can be surmised that a potential solution for educator adoption is a tool that (simultaneously) assists educators in working efficiently with a wide and flexible range of familiar data, addresses their felt needs while reducing workload, and can yield immediate, shareable benefits.

2.1 Addressing the entire 'data lifecycle' for educators

Existing approaches to learning analytics are predominantly dashboard-based, or mail-merge based. These are effective in addressing parts of the whole 'data lifecycle' that educators must manage through the course of a semester or year, but fail to address its entirety. Given the challenges highlighted above, this 'data lifecycle' in the context of personalizing feedback involves:

- Data collection – from online *and especially* offline learning and teaching environments. If feedback is to address where the learner is heading, how they are progressing, and what they can do to better achieve their goals (Hattie & Timperley, 2007), and if learning and assessment are occurring in disparate online *and* offline environments, then the right data needs to be available from both.

- Data curation – ensuring that all relevant data are brought into one place. The educator, with their understanding of the pedagogical and pastoral contexts of their course and the feedback needs of students, must be the one making the informed decisions about what data to curate in order to enable this feedback provision.
- Data manipulation and analysis – the 'raw' data may need to be transformed or otherwise manipulated before it can yield a useful representation of information, or be used to inform subsequent action.
- Actions enabled by the presence of data – providing feedback to students needs to occur in a timely way, account for individual student needs, and consider the classroom climate (Hattie & Timperley, 2007). Personalized feedback delivered by an electronic system (e.g. via email) helps to address this, but empowering educators (including tutors, coordinators, etc) with relevant data when interacting with students face-to-face and online is also important.

Only addressing parts of this lifecycle can lead to tools that serve one purpose well but fail to help educators who need a more comprehensive platform. For example, one tool might allow educators to import existing electronic data from spreadsheets and customize emails based on these data; this addresses part of data curation and action, but may miss crucial data from other environments or require educators to perform complex data manipulations before importing. Another tool might present to educators a set of appealing dashboards that represent pre-analyzed data; this may address part of data analysis but critically lacks educator ownership, educator choice of meaningful data, and data-enabled actions. We therefore contend that solutions that address parts of the whole data lifecycle will have limited utility and adoption because they still require users to perform (sometimes highly complex or time consuming) tasks outside the platform.

2.2 A potential solution: the Student Relationship Engagement System

The Student Relationship Engagement System (SRES; www.sres.io) is a unique learning analytics platform that has been developed to address these pressing needs and contextual challenges. In stark contrast to other LA approaches and tools, this platform, the Student Relationship Engagement System (SRES), gives precedence to teacher intelligence and small but meaningful data over predictive algorithms and big data. It enables educators to choose data that are important for their unique learning and teaching context (e.g. interim marks, attendance, tutor feedback, in-class participation grades, etc), and helps them to collect, collate, analyze, and make direct use of these data. Critically, educators can use the SRES to efficiently personalize learning support and feedback to students at scale by building simple rules to customize information that different students will receive via email, SMS, or a web page embedded into an LMS. For example, coordinators can use it to design a mobile-friendly SRES web app interface for tutors to record performance measures in class, and then build customized messages to be sent out to different students with suggestions for improvement based on these data. Teachers can also build interactive dashboards to visualize class trends and select sub-cohorts for follow-up. This puts educators in control of the whole data lifecycle, enabling them to obtain and use contextually-meaningful academic engagement and success data to foster relationships with, and belonging in, their students. These educator-student relationships are increasingly being recognized as playing a critical role in student engagement and persistence (Farr-Wharton, Charles, Keast, Woolcott, & Chamberlain, 2018), and a key factor in perceiving and receiving feedback (Hattie & Timperley, 2007).

3 AN IMPACTFUL BOTTOM-UP APPROACH?

Throughout the past six years, the SRES has been sustained and grown by a small team of committed educators and designers, without any funding. Primarily due to word-of-mouth recommendations and organic adoption, the SRES has increased its reach 15-fold since its inception in 2012 and is now used in over 20 departments for over 25,000 students at four Australian universities, with pilots commencing at several other institutions. Although the details of implementation varied, the common threads included: (i) faculty driving adoption from the bottom up, after seeing and/or experiencing the flexibility, impact, and efficiencies of the platform; (ii) the key role of 'third space' learning design staff in bridging the gap between the technology and the pedagogy; and (iii) agile or permissive information technology practices (Vigentini et al., 2017). Educators are customizing the SRES creatively to suit their students' contextual learning and feedback needs, and students are valuing the personalized feedback and support that educators provide through the platform. For example, over 70,000 pieces of personalized messages have been sent through the SRES at the three institutions over the last 18 months, and 96% of feedback-on-feedback from students (n=2250) indicate that these have been helpful.

REFERENCES

Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S., . . . Fisher, J. (2016). Student retention and learning analytics: a snapshot of Australian practices and a framework for advancement. In. Canberra, ACT: Australian Government Office for Learning and Teaching.

Farr-Wharton, B., Charles, M. B., Keast, R., Woolcott, G., & Chamberlain, D. (2018). Why lecturers still matter: the impact of lecturer-student exchange on student engagement and intention to leave university prematurely. *Higher Education, 75*(1), 167-185. doi:10.1007/s10734-017-0190-5

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). *Setting learning analytics in context: overcoming the barriers to large-scale adoption.* Paper presented at the Fourth International Conference on Learning Analytics And Knowledge, Indianapolis.

Gunn, C., McDonald, J., Donald, C., Blumenstein, M., & Milne, J. (2016). *The missing link for learning from analytics.* Paper presented at the ASCILITE 2016, Adelaide.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Macfadyen, L. P., & Dawson, S. (2012). Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Journal of Educational Technology & Society, 15*(3), 149-163.

Vigentini, L., Kondo, E., Samnick, K., Liu, D. Y. T., King, D., & Bridgeman, A. J. (2017). *Recipes for institutional adoption of a teacher-driven learning analytics tool: case studies from three Australian universities.* Paper presented at the ASCILITE 2017.

West, D., Huijser, H., Lizzio, A., Toohey, D., Miles, C., Searle, B., & Bronnimann, J. (2015). Learning Analytics: Assisting Universities with Student Retention, Final Report (Part 1). In: Australian Government Office for Learning and Teaching.

Combining technology and human intelligence to provide feedback and learning support using *OnTask*

Lisa Lim
University of South Australia
lisa.lim@unisa.edu.au

Sheridan Gentili
University of South Australia
sheridan.gentili@unisa.edu.au

Abelardo Pardo
University of Sydney
abelardo.pardo@sydney.edu.au

Shane Dawson
University of South Australia
shane.dawson@unisa.edu.au

Dragan Gašević
University of Edinburgh
dragan.gasevic@ed.ac.uk

ABSTRACT: Learning analytics (LA) research has been criticised as being too reliant on generic log data and student information. In essence, the derived data and analyses ignore the specific learning context thereby limiting the actionable intelligence that can be provided to students. In response to this, there has been a call for more contextualised LA-based approaches to feedback and support. This paper presents *OnTask*, an LA-based, technologically-mediated feedback system that allows course instructors to collate information on their students' learning and to send out personalised feedback. An important feature of this system is that it incorporates a course instructor's pedagogical knowledge of their course to select instructionally relevant indicators of engagement, to better direct students in their learning strategies. This paper discusses how the *OnTask* is aligned with key findings from research on feedback research, and describes its deployment in a large health-related course with a diverse student cohort. Preliminary findings from focus groups with students indicates that using *OnTask* for personalising feedback did promote greater self-regulation, and that this effect flowed over to the students learning in other courses. Further considerations for implementing technologically-mediated feedback are discussed.

Keywords: learning analytics, personalized feedback, large courses, higher education, self-regulated learning

1 BACKGROUND

Feedback is important for the enhancement of students' learning, however the massification of higher education and an increasingly diverse student population present significant challenges to course instructors to provide timely and actionable feedback (Khan & Pardo, 2016; Pardo, Poquet, Martinez-Maldonado, & Dawson, 2017). This is a critical issue as, students' satisfaction with the learning experience has been found to be positively associated with the amount and quality of feedback that students get (Jessop, El Hakim, & Gibbs, 2014). Yet, currently, research indicates that students are dissatisfied with the feedback they are receiving. One approach to address these challenges lies in the research associated with learning analytics (LA).

Research in Learning analytics (LA) has been oft criticised as being too reliant on generic learner log data and student demographics to generate early-warning systems (Gašević, Dawson, Rogers, & Gasevic, 2016). Though data-driven and therefore empirically based, employing such a generic approach across all courses is not sufficient to help students in terms of providing actionable intelligence (Wise, 2014) to improve their learning. Hence, more contextualised LA-based approaches to feedback and support are recommended by practitioners (Dawson, Jovanovic, Gašević, & Pardo, 2017; Liu, Bartimote-Aufflick, Pardo, & Bridgeman, 2017). As shown by Gašević et al (2016), the instructional conditions in a course cannot be neglected in the development of LA-based interventions. This paper introduces *OnTask*, an LA-based system that facilitates course instructors to provide "timely and personalized feedback and support, directly connected to each students' own learning data" (Liu, et al., 2017, p.147). This tool addresses two key challenges in providing feedback in higher education: 1) out of the vast amount of data available from the learning environment, the challenge of identifying and selecting indicators of engagement relevant to a specific course; 2) the challenge of scaling personal feedback to an increasingly diverse student population. This paper also describes its implementation in a large, health-related course in an Australian higher education institution.

2 OVERVIEW OF ONTASK

The *OnTask* system[1] was developed with the support of an Australian Government teaching and learning grant. *OnTask* is a software tool that collates information about students and their learning in a course, such as online engagement activity (from learning management system data), lesson attendance, and academic performance. The platform allows instructors to develop "if-then" rules to generate personalised messages to all students in their course. The same platform is then used to send out these emails. An important aspect of *OnTask* is that instructors choose the metrics that serve as indicators of engagement specific to the course, thereby providing more contextualised feedback and support. This is a process lacking in many generic LA-based systems (Liu et al., 2017).

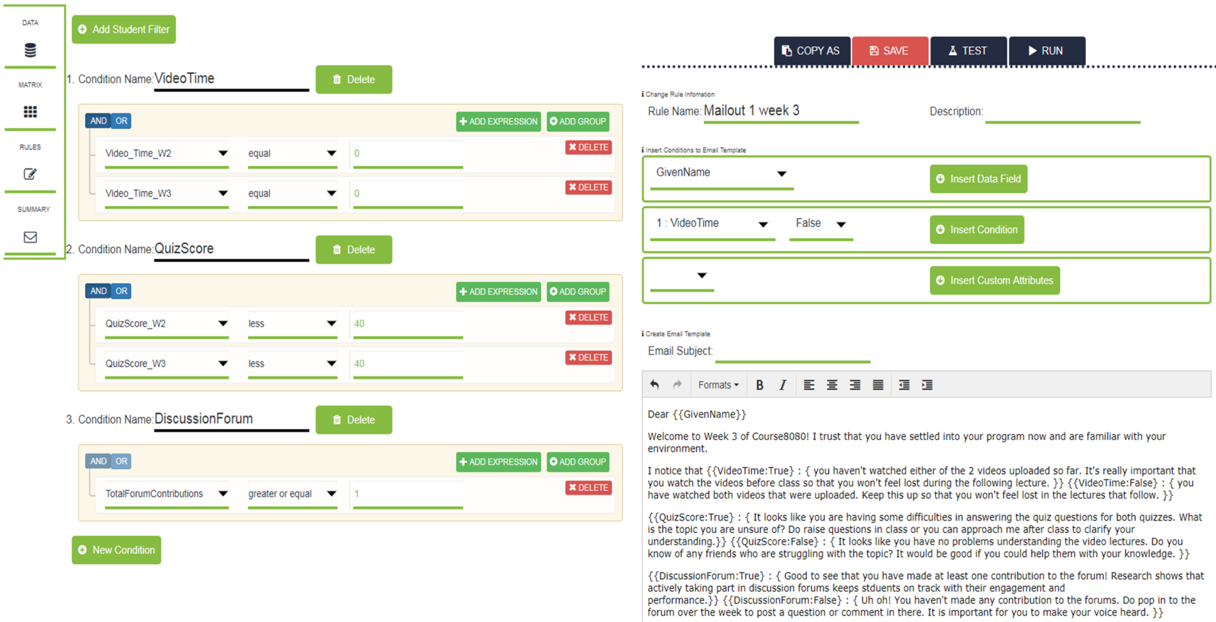
2.1 *OnTask* grounded in feedback research

The importance of feedback for learning has been well evidenced by John Hattie's (2014) widely cited meta-analysis of 150 variables affecting student achievement. Hattie demonstrates that studies examining the impact of feedback have a large effect size of .75, placing it within the top 10 influences

on achievement. However, the effects of feedback are not uniform. In particular, when it comes to impact on students' self-regulated learning, process feedback has been noted to be more effective than outcome feedback. Process feedback is more explicit about what the student needs to do in order to bridge a

[1] (https://www.ontasklearning.org)

learning gap and achieve a learning goal, while outcome feedback focuses on the result of the task, without further recommendation as to specific actions for improvement (Butler & Winne, 1995). Process



feedback could take the form of recommended actions, such as reading a specific article to strengthen understanding of a topic, or point out how certain learning strategies need to change. In order to support and develop students' self-regulated learning, and therefore their achievement, students need process feedback.

Figure 1: The *OnTask* interface showing "if-then" rule generation for personalised feedback

OnTask provides a platform for instructors to carry out effective feedback practices as identified by research. The following are specific principles of feedback enabled by *OnTask*:

1) *To develop mastery and deep learning, feedback should be aimed at the levels of process and self-regulation, rather than at the self* (Hattie & Timperley, 2007). *OnTask* collates data about online learning activities as well as from offline records such as attendance which can be uploaded into the system. In this way, instructors can provide feedback to their students about their learning progression and how they might adapt these to optimise learning in the course.

2) *Feedback should be given in a timely manner* (Shute, 2008). Feedback may be given either immediately after a learning activity or formative/ summative assessment task, or at a later point depending on the instructional intent. *OnTask* facilitates timely feedback by affording instructors opportunity to assign when the feedback (email) is to be disseminated. Moreover, instructors are also able to schedule specific and/or regular times for these messages to be automatically released, thereby reducing workload associated with having to provide regular feedback; this is a benefit especially for large courses.

3) *Feedback should be specific, or actionable* (Price, Handley, Millar, & O'Donovan, 2010). The giving of feedback should not be an administrative task to be checked off by instructors, but should facilitate students' adjustment of their learning towards the desired outcome or goal. Feedback is only effective when learners understand it and are willing and able to act on it (Shute, 2008). Because *OnTask* facilitates the collation of student engagement behaviours both online and offline, instructors can clearly define the activities around the engagement indicators which are necessary for closing the gap between students' current learning progress and the expected outcomes of the course. Knowledge about course demands and the effect of specified study behaviours on achievement helps students to self-regulate their learning more effectively (Ott, Robins, Haden, & Shephard, 2015).

4) *Feedback should have a positive tone, and prompt a dialogue with instructors about learning* (Nicol & Macfarlane - Dick, 2006). Feedback that is perceived as negative has been found to adversely impact students' reception, and therefore, their motivation to act on the feedback (Ryan & Henderson, 2017). Feedback should be seen as a process, involving students' sensemaking of the written content (Carless, 2016). To enhance this dialogic progress, instructors should follow up on feedback by consulting with their students. While *OnTask* uses technology to automatically combine digital traces and other evidence of student learning as feedback to students, the system also allows instructors to add a relational element to create personalised messages in ways to encourage student receptivity, e.g., through friendly or encouraging comments, and invitations for further face-to-face consultations.

3 USING ONTASK TO PROVIDE FEEDBACK AND LEARNING SUPPORT IN A HEALTH-RELATED COURSE

This section describes how *OnTask* was used in one higher education institution, focusing on the purpose and content of the feedback messages.

3.1 Course design and assessment

OnTask was trialled in two consecutive introductory biological sciences courses (Course A and Course B) with an enrolment of 242, at a public research-intensive university in Australia. The objective of Course A was for students to develop an understanding of the molecular and cellular aspects of biology, with topics such as basic cell physiology, macromolecules, pro- and eukaryotic cells. Prior knowledge in biology or chemistry was not required, which meant that the majority of students entering the courses had limited experience in these disciplines and therefore struggled with mastering the content. Lack of confidence and mastery impacted on both course A and subsequent courses (Course B). As an indication of diversity, student enrolment in these courses include 12 different programs ranging from health sciences to arts, and law.

Each course lasted for 13 weeks, after which students completed a final exam. The courses were based on a blended learning design. This involved online weekly activities combined with on-campus attendance in a 3-hour lecture and 1-hour tutorial. Over the semester, students were also required to attend 3 x 2-

hour work-shop sessions, and 7 x 3-hour Practical sessions. Students were requested to prepare for every Practical session by completing pre-lab activities online in the Learning Management System (LMS), as well as to reinforce their learning in the lectures by completing weekly online activities in an e-textbook, provided by McGraw-Hill *Connect*. Students were also informed at the start of the course that *Connect* was an important component of their learning, and that this platform would host two of their assessments (quizzes) contributing to their final course grade. Each course comprised four assessments: a mid-term quiz (10%), an end of term quiz (10%), Practical marks (25%), and final exam (55%).

3.2 Deployment of *OnTask*

A summary of how *OnTask* was deployed in Course A and Course B is provided in Table 1. The purpose and content of the mailouts reflected key events in the course: the first mailouts coincided with census period, with feedback on engagement indicators providing students with a review of their performance thus far in order to better inform their decision whether or not to withdraw from the course without penalty. Other than providing feedback, the emails were also used as a way to communicate with students about important events such as assessments, and to nudge students toward distributed practice rather than cramming for examinations (Seabrook, Brown, & Solity, 2005). In this respect, the *Connect* e-textbook activities were seen as especially important for reinforcing learning in the courses, and this was communicated with each mailout.

Table 1: Deployment of OnTask in Course A & Course B

Mailout	Week	Purpose	Engagement indicators used
		Course A	
1st	5	To remind about census date and key course assessments; to encourage students to engage with Connect; to convey importance about regular attendance	1. Connect registration status 2. Tutorial attendance 3. Workshop attendance
2nd	9	To provide feedback about quiz performance; to encourage students to engage with Connect	1. Mid-term quiz marks 2. Connect registration status 3. Connect engagement
		Course B	
1st	4	To remind about census date and key course assessments; to encourage students to engage with LMS & Connect; to convey importance about regular attendance	1. Connect registration status (Yes/No) 2. Tutorial attendance 3. Workshop attendance
2nd	6	To provide feedback on: engagement with LMS pre-lab activities; engagement with Connect; practical marks	1. Pre-lab activities completed (Yes/No) 2. Moodle activity (Not logged in at all) 3. Connect engagement (<50%) 4. Current Practical mark

3rd	8	To provide feedback on: engagement with LMS pre-lab activities; practical marks; engagement with Connect	1. Pre-lab activities completed 2. Connect engagement 3. Current Practical mark 4. Connect registration

3.3 Student perceptions

Preliminary findings from focus group discussions with students at the end of the course found that students appreciated the feedback support they had received through *OnTask*. Comments indicating that students' self-regulated learning was fostered with this approach are as follows:

> "It helps me to validate where I am; do I need to freak out right now?"

> "…gives you a nudge- Stop procrastinating and playing games!"

Students also appreciated the human element in the feedback, with a comment that *"The wording makes you want to do it. Like an encouragement"*. Particularly noteworthy was a comment that the feedback served as *"a reminder to study across the board",* indicating a flow-on effect to students' learning overall, not just in the present course of study.

4 NEXT STEPS

OnTask is currently being piloted in other higher education institutions in Australia and elsewhere. Other research reporting on the pilot uses of *OnTask* suggest that this approach to technology-mediated feedback combined with human intelligence has a positive impact on students' achievement (Fewster-Young, Chiera, & Schultz, 2017; Pardo, Jovanovic, Gasevic, & Dawson, 2017). Thus far it appears that students rely on the external feedback to calibrate their self-regulated learning. However, researchers (e.g., Orsmond & Merry, 2013) have highlighted that feedback should also foster the development of students' self-assessment practices. In view of this, implementation of any student-facing LA-based feedback system should consider the extent to which students are using the feedback as a crutch or as a scaffold, i.e., students should not be over-reliant on the feedback, but rather learn from the feedback how to accurately monitor their own learning process so that this instructional scaffolding can be gradually removed.

REFERENCES

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245-281.

Carless, D. (2016). Feedback as dialogue. In *Encyclopedia of Educational Philosophy and Theory, 1-6.* Accessed http://link.springer.com/referenceworkentry/10.1007/978-981-287-532-7_389-1.

Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017). From prediction to impact: Evaluation of a learning analytics retention program. In *Proceedings of the 7th International Conference on Learning Analytics and Knowledge*. ACM.

Fewster-Young, N., Chiera, B., & Schultz, L. (2017). *A personalised feedback mechanism in quantitative service courses to increase engagement and improve performance*. Paper presented at the

Australian Learning Analytics Summer Institute, ALASI 2017 (29 Nov - 1 Dec), The University of Queensland, Brisbane, Australia.

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*(Supplement C), 68-84. doi:https://doi.org/10.1016/j.iheduc.2015.10.002

Hattie, J. (2014). *Visible learning for teachers: Maximizing impact on learning*. Florence, GB: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi:10.3102/003465430298487

Jessop, T., El Hakim, Y., & Gibbs, G. (2014). The whole is greater than the sum of its parts: a large-scale study of students' learning in response to different programme assessment patterns. *Assessment & Evaluation in Higher Education, 39*(1), 73-88. doi:10.1080/02602938.2013.792108

Khan, I., & Pardo, A. (2016). Data2U: Scalable real time student feedback in active learning environments. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 249-253.

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-driven personalization of student learning support in higher education. In A. Peña-Ayala (Ed). *Learning Analytics: Fundaments, Applications, and Trends*, 143-169.

Nicol, D. J., & Macfarlane ‐ Dick, D. (2006). Formative assessment and self ‐ regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218. doi:10.1080/03075070600572090

Orsmond, P., & Merry, S. (2013). The importance of self-assessment in students' use of tutors' feedback: A qualitative study of high and non-high achieving biology undergraduates. *Assessment & Evaluation in Higher Education, 38*(6), 737-753. doi:10.1080/02602938.2012.697868

Ott, C., Robins, A., Haden, P., & Shephard, K. (2015). Illustrating performance indicators and course characteristics to support students' self-regulated learning in CS1. *Computer Science Education, 25*(2), 174-198. doi:10.1080/08993408.2015.1033129

Pardo, A., Jovanovic, J., Gasevic, D., & Dawson, S. (2017). Using learning analytics to scale the provision of personalised feedback. *British Journal Of Educational Technology (under review after minor revisions)*.

Pardo, A., Poquet, O., Martinez-Maldonado, R., & Dawson, S. (2017). *Provision of data-driven student feedback in LA and EDM*. Paper presented at the Handbook of Learning Analytics.

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: all that effort, but what is the effect? *Assessment & Evaluation in Higher Education, 35*(3), 277-289.

Ryan, T., & Henderson, M. (2017). Feeling feedback: students' emotional responses to educator feedback. *Assessment & Evaluation in Higher Education*, 1-13. doi:10.1080/02602938.2017.1416456

Seabrook, R., Brown, G. D. A., & Solity, J. E. (2005). Distributed and massed practice: from laboratory to classroom. *Applied Cognitive Psychology, 19*(1), 107-122. doi:10.1002/acp.1066

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research, 78*(1), 153-189. doi:10.3102/0034654307313795

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge*. doi:10.1145/2567574.2567588

Juggling system and data streams to personalize feedback: the case of a large first year course

Tania Bucic, Lorenzo Vigentini, Mark. E. King
UNSW Sydney, Australia
{t.bucic},{l.vigentini},{mark.king}@unsw.edu.au

ABSTRACT: This paper presents a large Business School foundation course developed as part of the Inspired Learning Initiative[1] at UNSW Sydney. Aiming to improve the student learning experience, the course design integrates several tools and technologies that optimize the way students work together, consume disciplinary content, and apply knowledge in constructively aligned assessments. The tools include case-based resources, external adaptive readings and self-assessment (from McGraw-Hill), team optimization (TMGrouper) and a newly developed tool to support the administration and management of communication with students (SRES). The results were positive, with two separate cohorts of students over two semesters reporting significantly higher satisfaction levels, higher overall grades, and lower failure rates than in previous years. This paper presents the design process and issues, and the overall effects on the student experience.

Keywords: personalization, feedback, e-learning, course design

1 INTRODUCTION

Educators routinely create courses with the assumption that students entering university are willing learners, eager to cultivate knowledge and acquire skills. Research suggests that students instead devote their focus on daily life management (Clydesdale, 2008, p.2). A recent 10-year review of the first-year experience in Australia (Baik, Naylor, & Arkoudis, 2015) showed that one-third of students experience a challenging and anxious journey because of the unknown and unclear expectations. Students consider dropping out for many complex and often interrelated reasons (Baik et al., 2015; Douglas, Douglas, McClelland, & Davies, 2015; Krause, Hartley, James, & McInnis, 2005; Pascarella, 2005; Tinto, 1987; Yorke, 1999). Evidence from business school students also reveals that academic motivation may not support educator assumptions. For example, more than 20% report that they are at university for social purposes and rate academic activities including studying, attending class, and homework among their lowest priorities (Krane & Cottreau, 1998). More than 40% are bored in class, and fewer than 35% study for a minimum of six hours each week, preferring instead to spend time working in gainful employment (Smart, Kelley, & Conant, 1999; Smart, Tomkovick, Jones, & Menon, 1999). We may therefore infer that business students are generally ill-prepared for academic success (Nonis, Hudson, Philhours, & Teng, 2005). Further, marketing students appear most likely to face academic difficulties compared to other disciplines

[1] The Inspired Learning Initiative (ILI) is a strategic program of work within the Portfolio of the Pro Vice Chancellor Education (PVCE) for an initial 5 years. This investment will enable the PVCE, in partnership with faculties, to supplement their existing resources in support of the educational development priorities of the UNSW 2025 strategy.

(Paladino, 2008). Business school graduates are expected to be able to make decisions in complex settings. Educators therefore strive to create active learning opportunities that present opportunities for engaged and collaborative learning (Taylor, Hunter, Melton, & Goodwin, 2011). Indeed, engagement has been shown to be linked with individual learner goal orientations (Bucic & Robinson, 2017), and motivational concerns appear to be enduring among marketing students (Freeman & Greenacre, 2011). In this study we explore how structural and behavioural interventions might influence experience uplift.

At UNSW Sydney, our focus on the first-year experience departs from the traditional transactional model of education delivery. Instead, we deploy an innovative delivery model using a total activity system, which includes internal, external and end-user partners. This multi-component model features multiple value creators that focuses on student experience. Personalised learning pathways and communications are customized using learning analytics and iteratively inform learning design. The total activity system is designed to address critical concerns that are particularly salient for first year students, including faultlines in group work, ambiguity in communications, and assessment anxiety. By scaffolding students in personalised ways, we uplift educational experience and improve student performance. Our innovative approach is embedded in an institutional perspective best summarized by the 2025 Strategy and the Scientia Education Experience, which is grounded on four key domains: Communities; Feedback and Dialogue; Inspired Learning through Inspiring Teaching; and Being Digital.

Within the strategic AUD $55.5 millions Inspired Learning Initiative, more than 600 courses have been targeted for 'Digital Uplift' over a five-year period – allowing for a programmatic approach to appraise, redesign and reevaluate the curriculum. This paper presents the case of one such course, Marketing Fundamentals (MARK1012), a large first year undergraduate course. The selection of this course for uplift provided a timely opportunity to leverage on the ILI, building into the process an experimental approach to course redevelopment which focused on two elements: 1) multiple engaging avenues for learning (flexibility/adaptability of modes of learning/teaching); 2) enhancement of feedback through targeted, personalised communications.

2 THE COURSE

MARK1012 is a large foundation course delivered each term at UNSW Sydney. About 800 students from across disciplinary spectrums complete the course each semester. MARK1012 introduces the student to the major concepts and theories of marketing, reflecting the breadth and diversity of the discipline. The course provides insights into where marketing fits within an organisation, its contributions to business in general, and outlines frameworks supporting marketing activities and challenges in the ever-changing market place. It utilises a value-based approach that is essential in practice to solve real-life business problems. The key learning objectives of the course are: 1) Describe core marketing concepts; 2) Understand the notion of value creation, value delivery and value capture; 3) Make marketing-based decisions. The delivery format is structured around 3 contact hours per week (weekly 2-hour lecture and 1 hour tutorial).

Given the opportunity provided by the ILI, the course convener was able to leverage a partnership with the Pro Vice Chancellor (Education) to integrate and deploy several tools to create the necessary infrastructure for course uplift.

2.1 Objectives of the 'Digital Uplift'

After an initial review of the course taught up to 2016, three main objectives were identified to drive the project:

1. A need to deliver a high impact, personalised experience through (i) customised feedback and (ii) targeted communication;
2. Enhancing logistics by streamlining the tutor-led activities and reporting. This includes enabling tutors to provide effective feedback via an integrating face-to-face, in-class assessment and asynchronous learning activities (online via Moodle, with MHCampus);
3. Foster student self-directed learning and autonomy by leveraging on the rich set of data generated in the course and by leveraging continuous feedback (e.g., prompts to stretch, revisit/review, commence readings/undertake self-tests, etc.)
4. At a later stage, reconsider the way in which content is delivered and assessment implemented (this includes piloting of different types of assessment, different deployment of feedback and the structure of teaching contact time)

3 TOOLS USED IN THE COURSE AND THE IMPLEMENTATION

Moodle™ is the main learning management system used at UNSW Sydney and provides standardised support for all courses. Academics are encouraged to use Moodle to offer a consistent look-and-feel across UNSW courses, and also to provide operational efficiency by hosting all course related resources, activities, support and evaluation tools. Moodle also offers a flexible platform permitting integration of external tools via LTI (Learning Tools Interoperability) and enables a relatively seamless experience for students. In the case of MARK1012, three additional tools were selected to augment the standard provision, each with specific pedagogical functions: TM Grouper suite, SRES and MHCampus from McGraw-Hill.

3.1 TMGrouper: Optimising team work and collaboration

The TMGrouper tool[2] is currently being developed and maintained in the portfolio of the PVCEas part of the PVCE Students as Partners initiative (one component of the ILI). The tool works to customise the formation of student teams based on instructor-set parameters and individual performance characteristics. For MARK1012, maximum heterogeneity between team members was the desired criterion. The tool extends work carried out at the University of Glasgow (Vigentini & D'Angelo, 2013) and

[2] Team grouper at https://grouper.teaching.unsw.edu.au

at the University of Melbourne (Bergey & King, 2014) showing that individual performance characteristics can be used for team formation in university courses, leading to improvement in effectiveness and, depending on the task, the degree of diversity can improve learning outcomes.

3.2 SRES: Improving communication via personalised feedback

SRES[3] is tool originally developed at the University of Sydney with the aim of improving relationship and engagement between teachers and students (Liu, Bartimote-Aufflick, Pardo, & Bridgeman, 2017). The underlying principle is that the efficient delivery of timely, personalised and actionable individual student feedback throughout the course, improves the learning experience and maintains a sense of connectedness to the course. The institutional implementation of the tool across multiple Australian universities shows evidence of the effectiveness of the model (Liu et al., 2017; Pardo, 2017; Vigentini et al., 2017).

3.3 MHCampus: adaptive learning via self- and summative assessment

MHCampus (including Connect and the LearnSmart platform)[4] integrates all of the digital products offered by McGraw-Hill Education with UNSW's LMS for quick and easy access to a suite of content and learning tools usually packaged with adopted textbooks. This integration allows to us to: seamlessly build an effective digital course grounded in an outcomes-based curriculum; automatically enrol students; design adaptive assessment (enhancing feedback provided to students); and provide an analytical view of student engagement, assessment and asynchronous interactivity that can be used to create a responsive teaching environment grounded on feedback and dialogue. The reports generated from this system allow for calibration of holistic student learning across face-to-face and blended formats. The *MHCampus* tool is positioned in the course ecosystem as a proof of concept adding delivery flexibility for the course. After its initial implementation, more courses will be added to the pilot in 2018 to support UNSW's diverse portfolio of programs. The flexibility and diagnostic superiority offered by *Connect* are currently unmatched by other products and are critical for developing adaptive capabilities required for future-proofing the educational experience.

3.4 Integrating tools

In the first iteration of the new-look course, a considerable amount of manual labour was required to integrate the range of tools and streamline the data flow between systems (see Figure 1). From top left to bottom right, student records for enrolment and class choice are automatically fed daily into Moodle. The LMS is connected to the McGraw-Hill products via LTI. The TM suite (including Profiler, Reporter and Grouper) is also connected to the LMS via LTI enabling a seamless experience for students. Finally, tutors

---

[3] SRES more information at https://goo.gl/DHAARp

[4] Connect is a discipline-specific platform that offer personalized study plans including unlimited practice questions, interactive learning aids, multimedia learning aids, adaptive follow up assignments, algorithmically generated values, and specific feedback on wrong answers, as well as a gradebook for learning analytics

log interactions with the students including attendance, class participation and assessment marks. Traditionally these data would be stored in individual spreadsheets collated by the course convener, but were automated via the SRES entry forms. For all other data sources an export step is required to bring the data into SRES. SRES is positioned as a conduit, providing an opportunity to collate all the course data in one place. As a result, the teacher is able to generate personalised messages to students based on their individual data points. The system also enables a quick-view the data for each student, each class, and the whole cohort, which would otherwise require a considerable time to prepare.

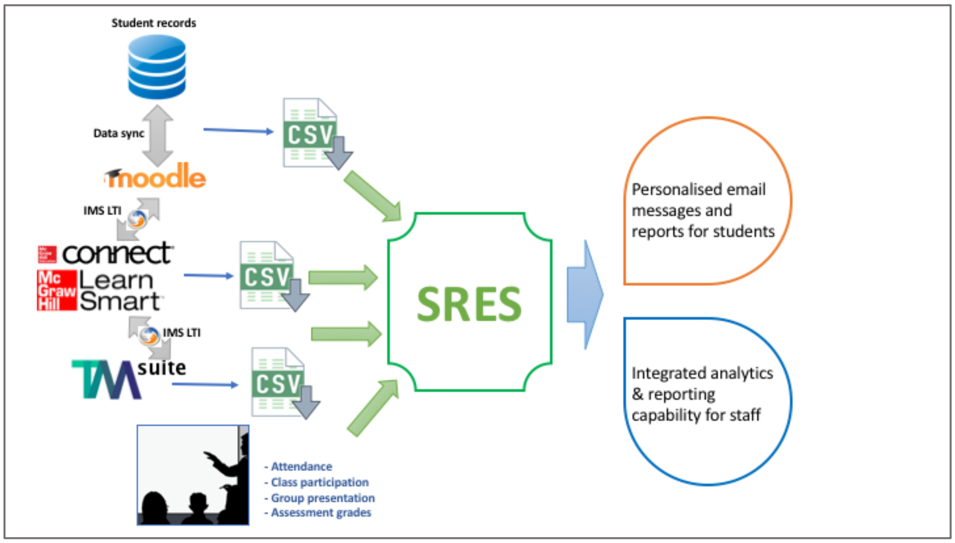


Figure 1: Overview of the system integration and data flow

3.5 Details about the cohorts and the 'experiment' (i.e. level of nudging)

In this paper, we present the MARK1012 results across two consecutive teaching periods. The course remained consistent across offerings with the only changes being the way feedback via personalised messages was provided and teacher presence. Table 1 provides an overview of the course structure and the timing of personalised messages during the course. Assessment components include two quizzes, a final report, student in-class participation, and a group-based 'case leadership' presentation.

Table 1: Overview of the course structure and feedback points.

	Topic	deadlines	SRES S1 feedback	SRES S2 feedback
Week0			Message1	Message1
Week1	Assessing the marketplace: Marketing Essentials		Message2	
Week2	Assessing the marketplace: Analysing the Marketing Environment		Message3	
Week3	Understanding and Targeting the Market: Consumer Behaviour			
Week4	Understanding and Targeting the Market: Segmentation, Targeting and Positioning	Quiz 1		
Week5	Understanding and Targeting the Market: Marketing Research		Message4	Message2
Week6	Value Creation: Product and Branding Decisions			

Week7	Value Creation: Developing New Products			
Week8	Value Creation: Services			
Week9	Value Capture and Delivery: Pricing			
Week10	Value Capture and Delivery: Supply Chain			
Week11	Value Capture and Delivery: Pricing		Message5	Message3
Week12	Course Review	Quiz 2		
Week13		Report	Message6	

The messages were intended to encourage students to complete activities and/or provide detailed feedback about their engagement with course activities. The key difference between the two runs of the course was in the level of nudging and frequency of messages to students, which can be described as 'teacher presence'. It should be noted that other course admin messages were broadcasted in a similar way to all students using the Moodle Discussion (which automatically notifies students via email) in both semesters, but the personalized messages were sent specifically from SRES, directly to individual student email accounts.

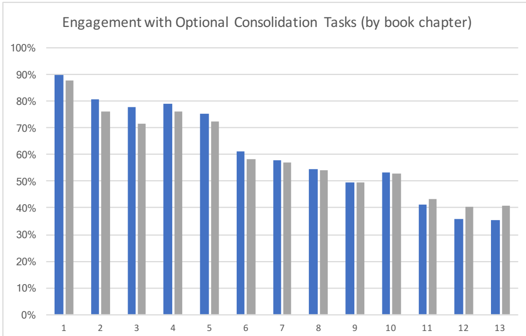
4 INITIAL RESULTS

Table 1 shows the large cohorts of students. Enrolled students are from diverse educational backgrounds, as demonstrated by the presence of 81 different degree programs from across all faculties. Although the majority of students are taking the course in their first or second year of study, between 15-20% of students take the course at a later stage. The uptake of MARK1012 may indicate the perceived usefulness of the course to a broader audience than traditional business courses. The table also provides an overview of the uptake of the digital platform (MGH and Learn Smart), with every student using the basic access provision (MGH) and a large majority (about 85% in the two semesters) choosing to (optionally) upgrade to include the Learn Smart adaptive ebook.

Table 1: Overview of key figures for the two semesters.

MARK1012	N students enrolled	Proportion of Y1/Y2	MGH platform uptake	LearnSmart uptake	Pass rate	High achievers	Satisfaction
S1_2017	795	82.4%	99.9%	86.67%	98.24%	84.91%	90.00%
S2_2017	860	85.1%	100.0%	84.42%	98.37%	91.74%	82.63%

The overall results show high student performance, with a pass rate (above 98%) and a growing number of students classed as high achievers with grades of 70% and above. The success of the course is also marked by a very high level of satisfaction with the quality of the course.



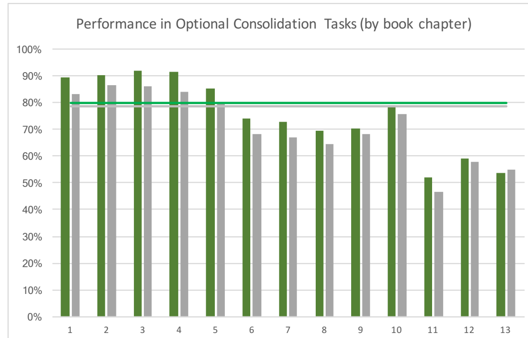


Figure 2. Distribution of completion and performance for the optional consolidation exercises by chapter (corresponding to weekly content). On the right, parallel bars represent the late stage quiz performance)

Interestingly, the effects of reduced teacher presence and reduction in nudging (halved compared to semester 1) had a negative effect on student satisfaction compared to the results in semester 1. Nevertheless, the results for s1 and also s2, show a marked improvement of an average of 10% compared to the same course in the previous two semesters (averaging 79.1%). In addition, the data confirm higher level of student achievement.

Further, examining interaction with the course tools, over 60% of students, on average, completed the optional weekly consolidation tasks in the MGH platform, which indicated the perceived implicit value of the platform for learning support. Modelling of the patterns of interaction in relation to the various performance outcomes is currently a work-in-progress.

Additional evidence from tutors' reports highlighted the positive effects of team optimisation. The two semesters reported the lowest level of complaints with teamwork ever received and tutors indicated their appreciation with the level of preparedness and in-class engagement across most students. Such a high level of engagement is also present in students' interactions with the systems and their engagement with the personalised messages (over 85% open rate across all messages).

5 DISCUSSION AND FUTURE DIRECTIONS

The evidence collated to date and the ongoing analytical work on the learning design as well as the non-cognitive elements of learning for students shows a noteworthy improvement in student performance and satisfaction. This work demonstrates the value of integrating several tools in the course. However, this resulted in a higher workload for staff involved in the organisation of the course (including both the academic lead and support staff). The piloting and adoption of tools such as TMGrouper has an additional overhead, but as it feeds directly into development, it means that the process can be improved over time. Similarly, there was a considerable amount of work required to extract, manipulate and organise the data used by SRES to personalise messages. From the perspective of scalability, further work is required to improve the level of integration and automation afforded by the assortment of tools.

Notwithstanding, the pilot described in this study has demonstrated the potential of such an approach to improve the quality of student experience and student performance. The positive and promising outcomes compel dialogue and investigation particularly concerning the impact of personalisation of stratified feedback at the course, school/discipline, and university levels. As demonstrated by the work done by Liu, Lim and Vigentini (2018), a growing number of case studies are becoming available to demonstrate the effectiveness of this approach and guide implementation of personalisation. The implementation model adopted for this MARK1012 also provided an extended opportunity to build new partnerships with capacity to create and deliver better educational experience.

REFERENCES

Baik, C., Naylor, R., & Arkoudis, S. (2015). *THE FIRST YEAR EXPERIENCE IN AUSTRALIAN UNIVERSITIES: FINDINGS FROM TWO DECADES, 1994-2014*. Melbourne: University of Melbourne. Retrieved from http://melbourne-cshe.unimelb.edu.au/__data/assets/pdf_file/0016/1513123/FYE-2014-FULL-report-FINAL-web.pdf

Bergey, P., & King, M. (2014). Team Machine: A Decision Support System for Team Formation. *Decision Sciences Journal of Innovative Education*, *12*(2), 109–130. https://doi.org/10.1111/dsji.12027

Bucic, T., & Robinson, L. J. (2017). Motivational engagement in the marketing classroom: individual goal orientations and class climate. *Journal of Strategic Marketing*, *25*(2), 164–177. https://doi.org/10.1080/0965254X.2016.1182579

Clydesdale, T. (2008). *The First Year Out: Understanding American Teens after High School*. University of Chicago Press.

Douglas, J. A., Douglas, A., McClelland, R. J., & Davies, J. (2015). Understanding student satisfaction and dissatisfaction: an interpretive study in the UK higher education context. *Studies in Higher Education*, *40*(2), 329–349. https://doi.org/10.1080/03075079.2013.842217

Freeman, L., & Greenacre, L. (2011). An Examination of Socially Destructive Behaviors in Group Work. *Journal of Marketing Education*, *33*(1), 5–17. https://doi.org/10.1177/0273475310389150

Krane, D., & Cottreau, A. (1998). Generation 2001: A Survey of the First College Graduating Class of the New Millennium. Study No. 718287. Retrieved from https://eric.ed.gov/?id=ED421945

Krause, K. L., Hartley, R., James, R., & McInnis, C. (2005). *The first year experience in Australian universities: Findings from a decade of national studies*. Centre for the Study of Higher Education, University of Melbourne Melbourne. Retrieved from http://www.cshe.unimelb.edu.au

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-Driven Personalization of Student Learning Support in Higher Education. In *Learning Analytics: Fundaments, Applications, and Trends* (pp. 143–169). Springer, Cham. https://doi.org/10.1007/978-3-319-52977-6_5

Nonis, S. A., Hudson, G. I., Philhours, M. J., & Teng, J. K. (2005). Changes in college student composition and implications for marketing education: revisiting predictors of academic success. *Journal of Business Research*, *58*(3), 321–329. https://doi.org/10.1016/j.jbusres.2003.06.001

Paladino, A. (2008). Creating an Interactive and Responsive Teaching Environment to Inspire Learning. *Journal of Marketing Education*, *30*(3), 185–188. https://doi.org/10.1177/0273475308318075

Pardo, A. (2017). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education*, *0*(0), 1–11. https://doi.org/10.1080/02602938.2017.1356905

Pascarella, E. T. (2005). *How college affects students*. San Francisco, Calif: Jossey-Bass.

Smart, D. T., Kelley, C. A., & Conant, J. S. (1999). Marketing Education in the Year 2000: Changes Observed and Challenges Anticipated. *Journal of Marketing Education*, *21*(3), 206–216. https://doi.org/10.1177/0273475399213006

Smart, D. T., Tomkovick, C., Jones, E., & Menon, A. (1999). Undergraduate Marketing Education in the 21st Century: Views from Three Institutions. *Marketing Education Review*, *9*(1), 1–9. https://doi.org/10.1080/10528008.1999.11488653

Taylor, S. A., Hunter, G. L., Melton, H., & Goodwin, S. A. (2011). Student Engagement and Marketing Classes. *Journal of Marketing Education*, *33*(1), 73–92. https://doi.org/10.1177/0273475310392542

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition.* ERIC.

Vigentini, L., & D'Angelo, A. (2013). Psychometric diversity in students team: what counts? In *Proceedings of the 18th Educational Learning Styles Information Network Conference*. Billund, Denmark: librix.eu.

Vigentini, L., Kondo, E., Samnick, K., Liu, D. Y., King, D., & Bridgeman, A. J. (2017). Recipes for institutional adoption of a teacher-driven learning analytics tool: Case studies from three Australian universities. In *ASCILITE Annual Conference Proceedings*. University of Southern Queensland. Retrieved from http://2017conference.ascilite.org/wp-content/uploads/2017/11/Full-VIGENTINI.pdf

Yorke, M. (1999). *Leaving Early: Undergraduate Non-completion in Higher Education*. Psychology Press.

Actionable Analytics at MyReviewers for Administrators, Instructors, Students, and Researchers[1,2]

Joseph Moxley
University of South Florida
mox@usf.edu

Spencer T. Bennington
University of South Florida
bennington1@mail.usf.edu

ABSTRACT: The goal of this paper is to provide an overview of existing and planned analytics at MyReviewers, an e-learning environment developed at the University of South Florida (USF) to expedite document review of texts, whether those texts are single-authored, collaboratively authored, alphabetic, or multimodal. Presently, MyReviewers serves as a site of collaborative knowledge making for students of multiple backgrounds in a variety of disciplines. This paper will explore the usefulness of analytics for multiple stakeholders. Ultimately, we conclude that it is the MyReviewers philosophy of "communal agency" and shared ownership.

Keywords: Writing Analytics, Self-Efficacy, Self-Regulation

1 OVERVIEW

MyReviewers was founded at USF (Tampa) in 2008 to support the General Education Program and to facilitate reporting to SACs (Southern Association of Colleges and Schools) accreditation. Funding was initially provided in 2008 by the General Education Council and Center for Teacher Enhancement. Between 2009 and 2012 support was provided by USF IT Funds. Beginning in 2013, support was provided by USF students enrolled in first-year composition (FYC) courses. That year also marked the introduction of MyReviewers into STEM classrooms like General Chemistry 1 and 2 and, eventually, Organic Chemistry. Since then, the program has been utilized in a variety of STEM courses so peers and instructors alike can provide formative feedback on genres ranging from lab reports to computer code.

MyReviewers features a variety of tools for students and instructors to use when evaluating a paper. Figure 1 illustrates many of these features visible on the main document markup page. Markup tools (1) include the "sticky note" and Community Comment™ functions as well as strikethrough and write-in capabilities. Instructors can easily toggle between students (2) with this menu bar or they can view a class gradebook for an overview of student progress. MyReviewers also features customizable rubrics (4) with dialog boxes (3) for rubric-specific commentary. These rubrics all have adjustable "milestones" (5) which allow instructors to assign specific qualities of writing to point values.

We believe the primary reason driving the success of Myreviewers is its foundation of transparency, dialogism, and communal agency. By developing the platform in response to the needs of various stakeholders in multiple contexts throughout the university, MyReviewers practices an incorporated approach to curriculum design. In short, MyReviewers provides stakeholders with a new way of conceptualizing and understanding learning analytics that makes the process of program development, professionalization, student learning, and research fully integrated and collaborative.

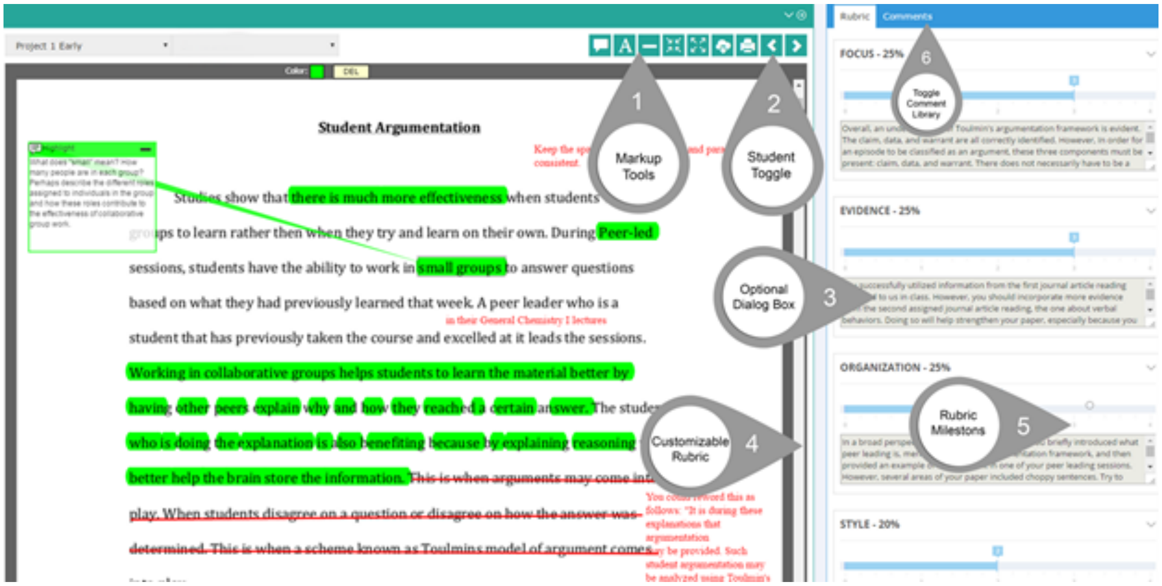


Figure 1: Markup Features in MyReviewers

2 MISSION OF MYREVIEWERS

MyReviewers endeavors to be more than a digital filing cabinet and more than a scoring platform. In fact, MyReviewers aims to be more than a simple tool for administrators, educators, students, and/or researchers. Instead, the MyReviewers platform seeks to embody an ethos of collaboration, one which embraces cutting-edge methodologies, encourages evidence-based revisions to pedagogies, and facilitates a better understanding of the intersections between writing instruction, computer science, quantitative data, and STEM education. In this way, MyReviewers is a living system, one that exemplifies the community of practice at USF from which it emerged. So far we have developed a vigorous interdisciplinary research community around the MyReviewers corpus, including an annual conference and an academic journal—*Journal of Writing Analytics*—published by Colorado State University Open Press. Studies under development include co-authorships with researchers from the Rand Foundation, ETS, Prairie State College, Montclair University, Carnegie Mellon University, Exeter University, Malmo University, and Tartu University.

Perhaps because it is a homegrown system, MyReviewers has not experienced much of the resistance to analytics reported in the literature. MyReviewers has been cultivated in response to the various immediate needs of individual stakeholders. When this is not the case, however, studies like Bailey and Garner (2010) reveal anxieties many educators possess with regard to making their feedback public to a corpus. This study explains how many instructors view big data in writing classrooms as just another administrative panopticon, one that aims to control and exert power over educators. When MyReviewers staff interviewed over 150 postsecondary faculty and administrators regarding their impressions of the platform as part of NSF I-Corps Grant #1636511, they found humanities faculty were particularly worried about the impact of analytics on disciplinary practices. This response aligns with some of the problematic attitudes outlined in Crossley and McNamara's (2017) discussion of incorporating technology in the classroom and is similar to the anxieties or resistance present in other studies like Herodotou et al's (2017) examination of 240 instructors at Open University, many faculty revealed resistance toward new technologies. In contrast, STEM faculty were much more ready to deploy tools like MyReviewers.

These STEM instructors believed tools like MyReviewers could help them save time and inform pedagogy with evidence. As stated in the EDUCAUSE Risks of Analytics Report, "The argument for analytics is that with large data sets, powerful analytics engines, and skillfully designed visualization techniques, we can use the experience of the past to create helpful models of our processes; we can even more effectively use real-time data and information to alert us to matters requiring our attention; and we can (in some cases) extrapolate to the future using predictive modeling and optimization techniques" (p. 12). This report demonstrates the time-saving features of analytics and describes how platforms like MyReviewers can streamline educational feedback. More importantly, perhaps, these receptive STEM faculty affirm the research of Bart Rienties and his team of data wranglers who found that the use of analytics is highly correlated with student learning and successful teaching (Nguyen, Rienties, Toetenel, Ferguson, & Whitelock, 2017). Where the Bailey and Garner (2010) study revealed doubts from instructors about the usefulness of feedback, Anson and Anson (2017) extended these doubts into research questions. Namely, the 2017 study examined the comparison between peer response and instructor feedback on a lexical and semantic level. Leveraging high quality feedback from over 50,000 peer responses provided by the MyReviewers corpus, Anson and Anson (2017) provided some of the first evidence that instructors who adopt the feedback practices of higher-order commenting can actually affect the peer response practices of their students. While this is just a first step, this study is a clear example of the kind of empirical evidence MyReviewers can provide to back up more theoretical claims with regard to notions of "generational shifts" in instructor response (Dixon and Moxley 2013). In other words, where once the field only had speculations or theories constructed from observation or experience, MyReviewers is providing a large data set to now test these hypotheses.

Of course, the use of writing analytics to effect change requires cooperation from multiple stakeholder groups. If instructors or administrators continue to resist technological advances in the classroom, as they have in the past decade, platforms like MyReviewers will remain largely ineffectual. If, instead, the multiple networks of academic, business, and private stakeholders work together to develop a "communal agency" as described in Vieregge et al.'s (2012) *Agency in the Age of Peer Production,* writing instruction as presently defined could be drastically improved for students of all disciplines. Moxley and Walkup's (2016) map of writing analytics describes just how collaborative this discipline must be in order to enact curricular change. But, as a variety of studies illustrate, this collaboration can yield success and opportunities that all participants previously thought were unattainable (Dixon and Moxley 2013, Moxley and Ross 2015, Donahue et al 2017). So the MyReviewers mission, then, is less about the tool itself, and more about the values it represents.

At USF, these values are integrated into a variety of networks. MyReviewers is utilized in First-Year Composition classrooms as well as STEM to explore how "new communication technologies . . . are empowering teachers, students, and Writing Program Administrators (WPAs) to radically transform...pedagogies—changing the roles of teachers and students, changing the content of our curricula, and changing our processes of composing and collaborating" (Moxley, p. 182, 2008). The MyReviewers platform itself is unique in that it is communally developed with the help of a wide constituency of administrators, instructors, graduate and undergraduate students, and professionals. This crowdsourcing allows for the platform to remain dynamic and on the cutting edge, constantly tuned to the needs of groups like the graduate teaching assistants in the FYC mentoring program who utilize it most. In this way, MyReviewers actually aids in the dissemination of knowledge between cohorts at USF, and allows for a constantly revised pedagogical approach not only to undergraduates, but to the rising professionals in the graduate programs. This is but one important illustration of how MyReviewers employs writing analytics not simply as a one-size-fits-all assessment, but as “context-specific, personalized, and designed to structure students’ opportunities to learn” (Moxley & Walkup, 2016). In the case of USF, these student learning opportunities include not only undergraduates seeking to improve their writing in a discipline, but graduate students seeking to improve their writing of instructor feedback.

Why is this mission so important? Moxley and Ross (2015) acknowledge that despite 50 years of research, writing studies is still struggling with many of the same questions posed decades before the information age. If MyReviewers can combine its technological power with the communal agency it attempts to embody, stakeholders will have a platform to engage with many of these longstanding questions, researchers will have the tools to find evidence to support answers, and educators can start shaping pedagogy based on data instead of lore (Langbehn et al. 2013; Moxley 2013). There are many features in the MyReviewers platform that allow for this.

For instructors, MyReviewers is designed to improve document and peer review processes. Features which benefit instructors include the ability to compare drafts to quickly ascertain whether students made effective revisions based on feedback. Instructors can also leave audio comments on texts or rubrics. This feature provides a wider range of disability access for users and increases grading speed for many instructors. Perhaps most impressive, however, is the collection of media-rich Community Comments™. These are pre-written feedback pages replete with video explanations, related articles, and sample exercises to check student comprehension all designed to reinforce lessons taught in class. The drag-and-drop function of these Community Comments™ also clarifies how to pinpoint theoretical concepts (like effective thesis statements, topic sentences, or paragraph transitions) in specific parts of a text. Community Comments™ expedite the review process for instructors and decreases the likelihood of grader fatigue (Dexter, 2014). This helps ensure that each student receives an equal measure of quality feedback and fair evaluations of their work.

This system of comments is also available to students for the process of peer review and helps train developing writers to better employ the language of appropriate critique. The MyReviewers platform provides multiple resources to help students understand the criticisms they receive from instructors and peers (four textbooks and a library of multimedia resources), features training modules to help students provide better feedback, and employs tools and workflows to make peer review faster.

MyReviewers is also designed to be a powerful assessment ecology for researchers. Students who opt in – that is, those who allow their texts and usage of the tool to be used by researchers for research purposes – complete writing background, self-efficacy, and self-regulation questions, which were developed by the writing program administrators at USF, MIT, Dartmouth, Penn, and NCSU for NSF Prime. The data gathered in the MyReviewers corpus allow for researchers to further investigate best practices and work with administrators to evaluate departmental efficacy.

3 MYREVIEWERS CORPUS

Since inception, over 50,126 students from over 12 universities have used MyReviewers to receive feedback from 1500 instructors and as well as conduct and view peer reviews. Presently, approximately 16,000 USF students use MyReviewers in a dozen undergraduate gateway courses: Composition 1; Composition 2; Professional Writing; Communication for Engineers; Technical Writing; General Chemistry 1 Labs; General Chemistry 2 Labs; Organic Chemistry 1 Labs; Organic Chemistry 2 Labs.

When students initially log on to MyReviewers, they are re-directed to a page that provides them with a Student Disclosure and Consent Form. Students are prompted to opt-in or opt-out of the research corpus before they gain access to the system. The consent form explains that no identifiable student data will ever be released to researchers or the public. Students are directed to a webpage that explains in detail why researchers are interested in developing the MyReviewers Corpus. These webpages explain that opting-in or out bears no relationship to grading or use of the system, and they explain how students can opt-out if they change their minds at any time. These pages define data and data usage policies. Students who decline to sign the student disclosure (opt-in) agreement are omitted from the data extracts available to researchers.

The USF part of the corpus includes over 187,000 student uploaded documents, 138,605 documents graded by instructors, and 209,103 peer review documents. During the 2016-2017 academic year, over 89,100 students papers were submitted—37,600 in the fall, 46,800 in the spring, and 4,600 in the summer. There was a 20% increase in student paper submissions between fall and spring. Over 74,300 peer reviews were also submitted—35,300 in the fall, 34,600 in the spring, and 4,300 in the summer. There was a 2% decrease in peer review submissions between fall and spring. Within this large body of documents, nearly 44,000 student papers have been opted-in to the extensive writing corpus—18,800 from the fall, 22,500 come from the spring, and 2,500 from the summer. Additionally, 23,400 peer reviews have been opted-in to the corpus—10,800 in the fall, 11,000 in the spring, and 1,500 in the summer. Researchers can use this corpus to analyze and identify factors that lead to strong student progress and program performance (Aull 2017; Anson and Anson 201).

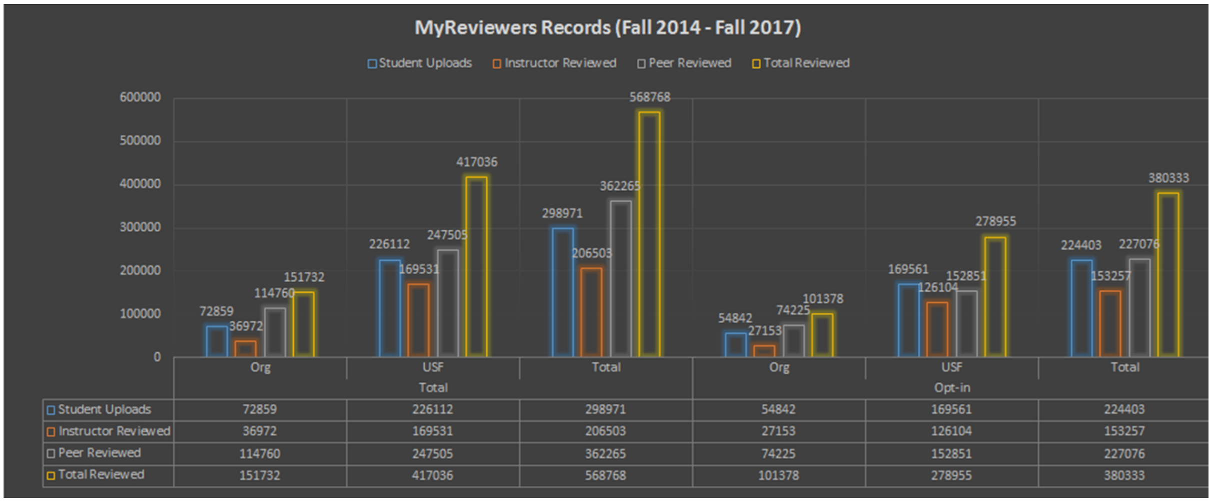


Figure 2: 2014- 2017 Reviewed Papers

Though this research corpus is sizable, the MyReviewers team is still working on increasing the number. Currently, opted-in student papers only account for 50% of all fall papers, 48% of all spring papers, and 55% of all summer papers, or roughly half of all papers submitted within the academic year. Meanwhile, opted-in peer reviews account for 31% of all fall reviews, 32% of all spring reviews, and 36% of all summer reviews, or roughly a third of all peer reviews submitted within the academic year. This means that, despite our best efforts, many students and professors are still on the fence about letting their work be researched, especially when it comes to peer reviews. We hope to demonstrate the value writing analytics can have on stakeholders to increase these opt-in numbers.

4 EXISTING ANALYTICS AT MYREVIEWERS

MyReviewers gathers data at a massive scale, and the current challenge is to ascertain how to visualize that data or ways to weigh that data in algorithms to advance students' opportunities to learn. As outlined below, the analytics provided to users depend on stakeholder role. Presently, the analytics in MyReviewers tend to provide descriptive information regarding statistical patterns, such as grades by course, grades by project, or grades by instructor. Eventually, the investigators hope to develop measures for competencies of interest to writing studies researchers, such as the development of students' cognitive, intrapersonal, or interpersonal competencies.

5 EXISTING ANALYTICS FOR ADMINISTRATORS

Administrator analytics are designed to meet the needs of administrators who are responsible for directing large gateway courses, such as composition, technical writing, or general chemistry courses (see https://www.youtube.com/watch?time_continue=24&v=3J-C95fthUw). These analytics are designed to aid administrators who have many faculty teaching multiple sections of a single course. The goal here is to further support and mentor graduate students and adjunct faculty. For instance, these analytics enable administrators to quickly view: Number of instructor-graded uploads, Number of completed peer reviews, Number of instructor-graded peer reviews, Number of completed revision plans, and Number of instructor-graded revision plans. Additionally, administrators may view: Grade Comparison: Compares the average grade the instructor gives by project; Grade Distribution: Compares the number of each letter grade the instructor gives; Word Count: Compares the average document word count by project; In-text Comments: Compares the number of in-text comments the instructor leaves; Community Comments™: Compares the number of Community Comments™ the instructor leaves. These analytics all provide valuable information to help ensure the administrator manages instructors most effectively for the purpose of protecting student interests. Furthermore, this allows administrators to target their efforts toward instructors who do not meet the average pace or quality of other instructors when it comes to grading effectively or expediently. From a programmatic standpoint, this leads to a higher concentration of well-trained instructors.

6 EXISTING ANALYTICS FOR INSTRUCTORS

As illustrated at https://www.youtube.com/watch?v=HWMdzckifZk, instructor analytics at MyReviewers are designed to help instructors identify patterns in student data, such as common critiques instructors are providing across drafts or data regarding an individual instructors' scoring patterns compared with other instructors teaching different sections of the same class. For instance, instructors may view: Grade Comparison: Compares the average grade the instructor gives by project; Grade Distribution: Compares the number of each letter grade the instructor gives; Word Count: Compares the average document word count by project; In-text Comments: Compares the number of in-text comments the instructor leaves; Community Comments™: Compares the number of Community Comments™ the instructor leaves.

This information provides immediate feedback for an instructor to help tailor further lessons toward the specific needs of their students. Just as the administrator can more effectively manage instructors, the instructor can more efficiently respond to the educational needs of individual classes.

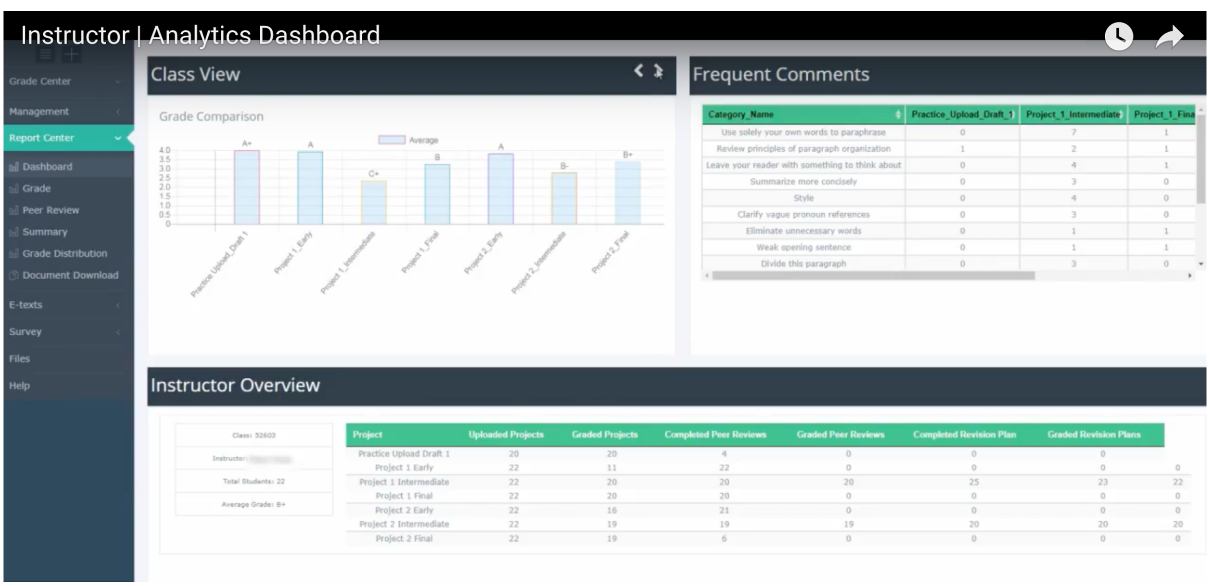


Figure 3: 2016- 2017 One of the Instructor Dashboards

7 EXISTING ANALYTICS FOR STUDENTS

Presently, the goal of the Student Analytics is to help students track critiques of their work over time, to compare their performance to the class mean and exemplars, and to reflect on their self-efficacy and self-regulation (see https://www.youtube.com/watch?v=v_P88EXg9Y8). For instance, sample reports include Grade Comparison: Compares the student's grades to the class average and exemplary students; Criteria Scores: Compares the student's rubric criteria scores by project to the class average and exemplary students; Document Word Count: Compares the student's document word count to the class average; Peer Review Comments Word Count: Compares the student's average comment word count to the class average and exemplary students; Helpfulness: Compares how peers rated the helpfulness of a student's feedback in the rate-the rater survey to the class average and exemplary students. For student writers, MyReviewers aggregates and visualizes critical comments in ways that promote and measure revision.

One of the most beneficial aspects of these student facing analytics is they provide real-time feedback at every graded juncture in a course. This encourages students to seek out instructor interventions more readily as they have a clearer idea of how they are performing in the class. If a student notices they are performing at a lower rate than the class average, they can investigate why that might be true by examining the patterns established over time from their scored drafts.

8 FUTURE ANALYTICS AT MYREVIEWERS

The MyReviewers development team seeks partners for research and development. We are eager to imagine new analytics, such as those that use sentiment analysis to categorize feedback, to advance students' opportunities to learn, and to help instructors save time while providing more useful, motivational feedback. The platform has made some progress, but it still exists in a stage where it provides mostly descriptive analytics in the form of different dashboards for multiple stakeholders. The investigators recognize this as a first step, however, and they are currently working toward ways in which these analytics can go beyond simple descriptions of data.

Future endeavors include utilizing machine learning and neural networking technologies to help students navigate some of the platform resources more effectively. For example, investigators received an NSF SBIR Phase 1 grant #1721749 for *Artificial Intelligence, Scientific Reasoning, and Formative Feedback: Structuring Success for STEM Students*. This research aims to build AI that can (1) identify scientific reasoning in student lab reports; (2) score the labs on a 1 to 3 scale; and (3) use writing analytics to suggest comments that hyperlink to an article, video, and *try it* exercises related to the comment. Thus far we have developed several machine learning methods, with Random Forest and Support Vector Machines being the top performers, to grade the lab reports. We have built machine learning models targeting two facets of writing: fluency and topic-specific vocabulary. The average accuracy of the models ranges from 60% to 100% per text section, which is supported by high levels of statistically significant correlations among fluency measures and scores for each section of the labs. This is but one of many steps the MyReviewers platform is taking into a new era of educational research and peer production. Future analytics may also be mapped onto the non-cognitive domains of learning and help inform administrators, educators, and students of ways in which competencies like creativity, flexibility, or determination might be further developed or practiced in the classroom.

Other research projects currently underway include studies involving student perception of the peer review process, investigations into the lexical and semantic characteristics of effective student writing, swalesian analyses of the genre conventions and rhetorical moves present in effective scientific writing, and explorations of alternative document markup strategies. Future researchers will also be interested to know that MyReviewers is currently developing a more interactive corpus archive for non-technical users. This will help enable more successful replication studies as well as a variety of future endeavors.

One truth remains: no future research can be successful without the buy-in of a diversity of interested stakeholders. In the past, platforms like MyReviewers have struggled with this because of various misconceptions or fears surrounding learning analytics. We believe that it is extremely important, now more than ever, to reimagine the theoretical structure of how analytics are integrated within educational systems. MyReviewers' success is due to its embodiment of peer production, collaboration, and communal agency. MyReviewers acts as a dialogical support system for administrators, teachers, students, and researchers, and it continues to develop in response to the needs of all stakeholders. This mission, above all else, is what MyReviewers aims to uphold.

NOTES

1. This research has been funded by the National Science Foundation: NSF DGE #1544239: Collaborative Research: The Role of Instructor and Peer Feedback in Improving the Cognitive, Interpersonal, and Intrapersonal Competencies of Student Writers in STEM Courses

2. Moxley has a financial interest in MyReviewers LLC, which licenses the MyReviewers software from USF. This interested has been viewed and managed by the University in accordance with its individual and institutional Conflict of Interest policies.

9 REFERENCES

Anson, I.G. & Anson, C.M. (2017). Assessing peer and instructor response to writing: A corpus analysis from an expert survey. *Assessing Writing, 33 (2017)*, 12-24.

Aull, L. L. (2017). Tools and Tech: A New Forum. *Assessing Writing*, *33*A2-A7.

Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education, 15,* 2, 187-198.

Crossley, S. A., & McNamara, D. S. (2017). *Adaptive educational technologies for literacy instruction. [electronic resource]*. New York, N.Y. : Routledge, 2017.

Dexter, E. S. (2014) The Effect of Fatigue or Boredom on Teachers Marks. *The Journal of Educational Research, 28,* 9, 664-668.

Dixon, Z., & Moxley, J. (2013). Everything is illuminated: What big data can tell us about teacher commentary. Assessing Writing, 18(4), 241–256.

Donahue, C., Elliot, N., Ross, V., & Moxley, J. (2017). At the intersection of writing program administration, the digital humanities, and STEM education: Corpus methods as a lens for reader response. Paper presented at the Modern Language Association Convention. Philadelphia, PA.

Herodotou, C., Rienties, B., Boroowa, A., Zdrahal, Z., Hlosta, M., & Naydenova, G. (2017). *Implementing predictive learning analytics on a large scale: The teacher's perspective*. Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC.

Langbehn, K., McIntyre, M., & Moxley, J. (2013). Re-mediating writing program assessment. Presented at the Digital Writing Assessment & Evaluation, Logan, UT: Computers and Composition Digital Press.

Moxley, J. (2008). Datagogies, Writing Spaces, and the Age of Peer Production. *Computers and Composition, 25,* 2, 182-202.

Moxley, J. (2013). Big data, learning analytics, and social assessment. Journal of Writing Assessment, 6(1)

Moxley, J., and Ross, V. (2015) Using digital tools to foster writing research and student success in STEM courses.

Moxley, J., & Walkup, K. (2016). Mapping writing analytics. In Proceedings of the 9th International Conference on Educational Data Mining. Raleigh, NC: EDM.

Nguyen, Q., Rienties, B., Toetenel, L., Ferguson, F., & Whitelock, D. (2017). Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates. *Computers in Human Behavior*, *76*(November 2017), 703–714.

Probabilistic Graphical Models as Personalised Feedback

Mary Loftus
National University of Ireland, Galway
mary.loftus@nuigalway.ie

Michael G. Madden
National University of Ireland, Galway
michael.madden@nuigalway.ie

ABSTRACT: Given appropriate tools and starter data, learners can gather data themselves to build personalised feedback models. Learners are often seen as passive objects of learning analytics (LA) initiatives and the opaque, black-box nature of many of LA systems reinforces this tendency. This research explores ways to involve students as active participants in shaping LA to their needs and encouraging them to be critical, data literate users of their own learning data. We have developed proof-of-concept, interactive, white-box probabilistic graphical models which students can interact with visually, and which are designed to give a sense of agency and an interconnected understanding of their learning activities. We are using qualitative methods to try to learn what other nodes would be impactful for students' understanding and that would prompt and support metacognitive and self-reflection activities. These models are being developed further, tested and validated. Finally in future phases, we will conduct quantitative and qualitative analyses of their impact on student learning and learners.

Keywords: Learning Analytics, Probabilistic Graphical Models, Bayesian Networks, Probability, Metacognition, Self-Reflection, Personalised Feedback, Data Literacy, Ethics

1 BACKGROUND

'The invention of the camera has changed not only what we see, but how we see it', observed John Berger in his ground-breaking BBC series 'Ways of Seeing' (1972) as he analysed the impact of mass media on society and its evolution from its roots in art. Today, the evolution of Learning Analytics is again changing what we see in learning; it is also changing how we see learning and it is beginning to change the nature of learning itself.

1.1 What is Education?

Biesta's call to define the purpose of education as 'the coming into presence of unique individual beings' (2015) serves as a useful starting point for this research. We must be clear about the purpose of education and the role of Learning Analytics in our learning spaces if we are not to blindly take on board the social and political constructs embedded in data analytics approaches from other fields like industry. We have an opportunity here to use data in our classrooms to see into new corners and foster critical thinking. A key part of the education process is the 'individuation' or 'subjectification' of each human being – 'the process of becoming a subject' (Biesta, 2015). This is the ontological starting point of this research along with Paulo Freire's (1970) emphasis on the student as an agent of praxis in their learning environment.

1.2 Can students generate learning feedback themselves – using a critical approach to learning analytics?

One of the key questions driving this research is whether we can use learning analytics as a means of praxis and empowering students to use data to 'read the world' around them (Freire, 1970). We must also be cognisant, and help students be cognisant, that data is political and can replicate societal bias (boyd, 2017). Prinsloo & Slade challenge us to use analytics in ways that:

> *'decrease student vulnerability, increase their agency, and empower them as participants in learning analytics — moving them from quantified data objects to qualified and qualifying selves.* (Prinsloo & Slade, 2016)

There is an opportunity here for students to see that they themselves can be cast as objects in a system of systems, that their data is the commodity of corporations and others, that inequality is reflected in these data-driven systems, and that they must become proactive in managing their digital footprint, their data and their privacy in order to thrive as citizens of the future.

Can we scaffold learners in building their own learning analytics models using simple but powerful visual representations of their own data that they can use to reason about their own learning and reflect on how they learn?

1.3 Research Values: A Critical Approach and Sound Ethical Grounding:

- Grounded in the Student perspective - with ethics front and centre
- Encourages and supports students as owners and users of their learning data
- Uses Machine Learning techniques emphasising white-box, relational modelling
- Builds capacity for Data literacy with students
- Co-designs new learning analytics models that give students a more cohesive insight into their learning activities, goals and dispositions

This research builds on the work of modelling for and by students – using white-box approaches – Bayesian Networks in particular, a subclass of Probabilistic Graphical Models.

1.4 Probabilistic Graphical Models

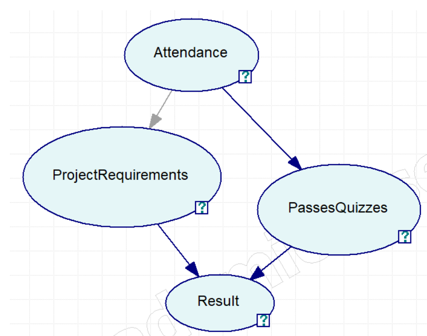


Figure 1- Bayesian Network Model

There are many modelling approaches (Chrysafiadi & Virvou, 2013) but not all are accessible to the student themselves and not all lend themselves to effective reasoning approaches. Probabilistic Graphical Modelling is a branch of machine learning that studies how to use probability distributions to describe the world and to make useful predictions about it, combining Bayesian probability and graph theory. Bayesian Networks (Fig 1) are simple constructs in some ways but have been proven to be a powerful tool in student modelling (Millán, Loboda, & Pérez-de-la-Cruz, 2010). In education settings, Bayesian networks have seen specific application to Intelligent Tutoring Systems (ITS) in Bayesian Knowledge Tracing (Corbett & Anderson, 1994; Baker, Corbett, & Aleven, 2008), metacognitive modelling (Conati, Gertner, & Vanlehn, 2002), student assessment (Martin & VanLehn, 1995), multi-agent tutoring systems (Zapata-Rivera & Greer, 2004). Hidden Markov Models are another class of PGM that have been used effectively to model student learning (Piech, Sahami, Koller, Cooper, & Blikstein, 2012).

1.5 What kind of feedback can be generated?

This approach to generating feedback has the potential to be proactive and student-led. It is formative in nature, as students will be getting feedback right at the start of a learning experience, based on data from previous cohorts and as regularly as they wish from then onwards. One of the criticisms of learning analytics (Lodge & Lewis, 2012) is that it can give a reductive and behaviourist view of the complex process of learning – and perhaps drive the wrong behaviours. This needs to be part an ongoing conversation we have with students, alerting them to the dangers of just 'gaming the system' and the potential for a 'Hawthorne Effect' (Adair, 1984) and emphasising the potential for generating new insights and self-reflection when the student engages positively with the system.

2 RESEARCH METHODOLOGY, WORK COMPLETED & NEXT STEPS

Course data has been gathered with students in a 1yr 'Internet of Things' module and has been parsed into interactive Bayesian Networks. These will be shared with the next cohort of students and a Grounded Theory study will be completed to build a theory of what models students find useful and more importantly what other data models would be perceived as helpful and insightful for them as they evolve as learners. New models from this qualitative work with students will be built and presented back to students for another iteration of quantitative and qualitative analysis.

2.1 Ethics Approval Granted & Initial Literature Review Completed

In January 2017, the NUI Galway Research Ethics Committee considered an application for ethics approval for this work and approval was granted on the 1st Feb 2017. An initial literature review has been completed with a bibliography of over 300 hundred papers.

2.2 Data Gathered so far

Data has been gathered in a first year 'Internet of Things' module – which the authors teach as part of a BSc in Computing programme and this includes:

- Students' activity in Moodle VLE
- Attendance data
- Students' activity in Github & Trello
- Quiz result
- Final Grade

2.3 Data Pre-processing & Initial Modelling Completed

We completed work on collating and discretising this data and building initial Bayesian Networks models in Sep/Oct 2017. The example model in the next section (Figures 2 & 3) shows how students can interact with course variables and see expected correlations with course performance, based on the performance of a past cohort. Students will be able to condition these variables on their own actual or expected performance levels. Eventually, either directly or indirectly, they will be able to add new nodes that they believe are important to their educational experience. While these current proof-of-concept models are an interesting starting point, we believe their real impact will only become fully apparent when we start to model on nodes that emerge from our qualitative work with students on identifying their own learning data prorities.

2.4 Validating the models

Work has begun on validating these models for predictive accuracy – using training, test and hold-out datasets. We also want to evaluate how well they will generalise to our new cohort of students. We are also evaluating ways to subjectively validate the models' impact on student self-reflection.

2.5 Output: Initial proof-of-concept Bayesian Network models designed to have students interact with the interconnected performance variables in their course

In Figures 2 & 3, students can see a view of their future on Day 1 of their course. This model was built on data from the previous cohort's performance but allows current students to condition on their own actual or forecasted performance. The intention here is to set students up for success from an early point in their course, giving them a data-informed understanding of the impact of their learning activities.

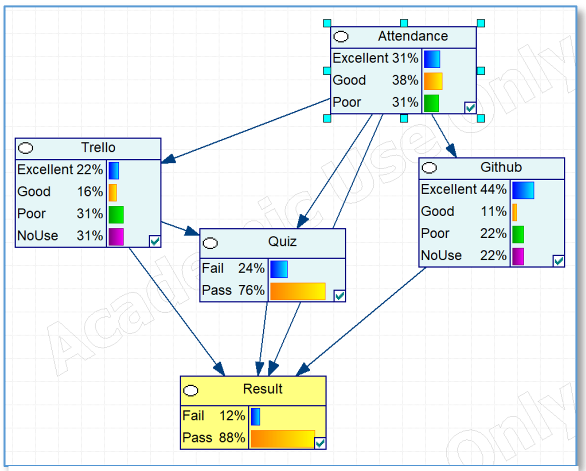


Figure 2 - A Bayesian Network model showing students some of the interconnected performance variables in their course

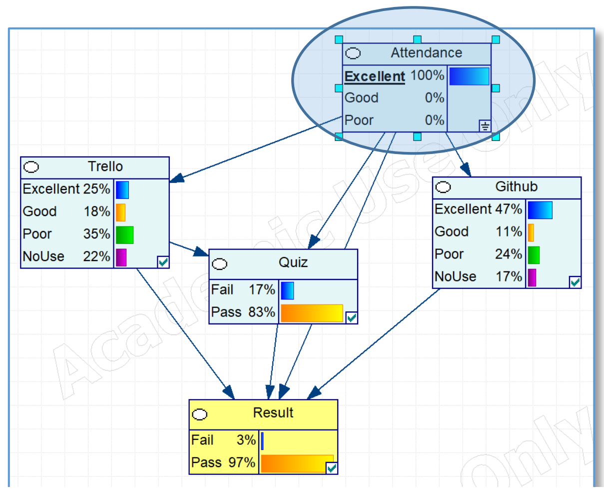


Figure 3 - Students can condition on any variable – e.g. 'Excellent Attendance' propagates around the network to show impact on other variables and future performance

2.6 Presentation of Initial Findings

Initial findings from this work were presented at a seminar entitled ***'Critical Learning Analytics'*** hosted by the *Society for Research in Higher Education (SRHE)* (Loftus, 2017) and at Ireland's Edtech conference (Loftus, 2017b). An earlier version of the work won 'Best Interdisciplinary Paper' at an NUI Galway/University of Limerick postgraduate Research Day. (Loftus, 2017a)

2.7 Next Steps

2.7.1 Building & validating more substantial Bayesian Networks - the initial proof-of-concept models are currently being validated, extended and tested for accuracy using hold-out datasets.

2.7.2 Qualitative Research – the voice of the student – we are conducting some qualitative research with students, using a Grounded Theory approach, to get student

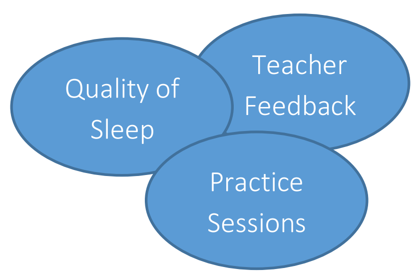


Figure 4 – Nodes Students Might Add?

perspectives on the data that might be usefully modelled to support student progression, critical thinking and class-based metacognitive activities.

2.7.3 Design of a tool to allow students to construct & interact with their own learning data models. The Connected Learning Analytics Toolkit is a candidate base platform for this work (Kitto et al., 2016).

2.7.4 Design of pedagogical materials to facilitate discussions with students, provide some basic data literacy skills and awareness of how data can be used in learning and elsewhere.

2.7.5 Quantitative/qualitative analysis of students' use of the new LA tool – this work will assess the impact of the research project on learning and learner outcomes.

REFERENCES

Adair, J. G. (1984). The Hawthorne effect: a reconsideration of the methodological artifact. *Journal of Applied Psychology*, *69*(2), 334.

Baker, R. S. J. d., Corbett, A. T., & Aleven, V. (2008). More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent Tutoring Systems* (Vol. 5091, pp. 406–415). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/10.1007/978-3-540-69132-7_44

Berger, J. (1972). *Ways of seeing* (Vol. 1). Penguin uK.

Biesta, G. J. J. (2015). *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Routledge.

boyd, danah. (2017, April 12). Toward Accountability. Retrieved 18 April 2017, from https://points.datasociety.net/toward-accountability-6096e38878f0

Chrysafiadi, K., & Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, *40*(11), 4715–4729. https://doi.org/10.1016/j.eswa.2013.02.007

Conati, C., Gertner, A., & Vanlehn, K. (2002). Using Bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, *12*(4), 371–417.

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, *4*(4), 253–278.

Freire, P. (1970). Pedagogy of the Oppressed. 1970. *New York: Continuum*, 125.

Kitto, K., Gašević, D., Siemens, G., Lynch, G., Bakharia, A., Lupton, M., … Dawson, S. (2016). The connected learning analytics toolkit (pp. 548–549). ACM Press. https://doi.org/10.1145/2883851.2883881

Lodge, J., & Lewis, M. (2012). Pigeon pecks and mouse clicks: Putting the learning back into learning analytics. *Future Challenges, Sustainable Futures. Proceedings Ascilite Wellington*, 560–564.

Loftus, M. (2017a, April). *Ways of Seeing Student Learning & Metacognition with Machine Learning and Learning Models*. EDTECH 17 - TEL in an Age of Supercomplexity - Challenges, Opportunities and Strategies. Retrieved from https://nuigulresearchday2017.wordpress.com/interdisciplinary-abstracts/

Loftus, M. (2017b, June). *Ways of Seeing Student Learning & Metacognition with Machine Learning and Learning Models*. EDTECH 17 - TEL in an Age of Supercomplexity - Challenges, Opportunities and Strategies. Retrieved from http://programme.exordo.com/edtech2017/delegates/presentation/137/

Loftus, M. (2017, October). *Ways of Seeing Learning Through Data*. Presented at the Critical Learning Analytics, Society for Research in Higher Education (SRHE). Retrieved from https://www.srhe.ac.uk/events/details.asp?EID=322

Martin, J., & VanLehn, K. (1995). Student assessment using Bayesian nets. *International Journal of Human-Computer Studies*, *42*(6), 575–591. https://doi.org/10.1006/ijhc.1995.1025

Millán, E., Loboda, T., & Pérez-de-la-Cruz, J. L. (2010). Bayesian networks for student model engineering. *Computers & Education*, *55*(4), 1663–1683. https://doi.org/10.1016/j.compedu.2010.07.010

Piech, C., Sahami, M., Koller, D., Cooper, S., & Blikstein, P. (2012). Modeling how students learn to program. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education* (pp. 153–160). ACM.

Prinsloo, P., & Slade, S. (2016). Student vulnerability, agency, and learning analytics: An exploration. *Journal of Learning Analytics*, 159–182. https://doi.org/10.18608/jla.2016.31.10

Zapata-Rivera, J.-D., & Greer, J. (2004). Inspectable Bayesian student modelling servers in multi-agent tutoring systems. *International Journal of Human-Computer Studies*, *61*(4), 535–563.

Personalizing Non-Cognitive Feedback for Learning and Skill Formation

Alfred Essa
McGraw-Hill Education
alfred.essa@mheducation.com

Ani Aghababyan
McGraw-Hill Education
ani.aghababyan@mheducation.com

ABSTRACT: An emerging body of empirical research indicates the importance of non-cognitive skills (N-skills) for a variety of outcomes, including wealth, health, and education. Among N-skills are confidence, motivation, self-regulation, grit, mindset, and procrastination.

Current approaches to the study of N-skills suffer from several limitations. First, N-skills are often considered singly, without a clear understanding of their relationship to each other. Second, N-skills are considered statically, as fixed attributes which don't change over time. Third, the term "non-cognitive" skills is a misnomer since N-skills often have an underlying cognitive component. Finally, N-skills are presented as if they do not interact with cognitive skills (C-skills).

In the presentation we first outline a theoretical econometric model of skill formation (derived from James Heckman and Flavio Cunha) as a framework for organizing disparate findings in psychology, education, and neuroscience. The model presents a dynamic view of skill formation which takes into account both C- and N-skills as well as their interactions.

Using the model as a starting point we provide several analytical examples for studying the dynamic characteristics of N-skills:

- The concept of information entropy is used to illustrate the variability of "confidence" over time.
- A "procrastination index" is formulated as a way of studying habits and the stability of preferences.
- Data mining is applied to identify learner profiles which are composites of C-skills and N-skills

Based on the econometric model and analytic examples we argue that personalized feedback can be viewed simultaneously as interventions and investments. Just-in-time interventions, cognitive and non-cognitive, can facilitate in the moment learning. But consistent interventions over time can also be viewed as investments which assist learners with skill formation and the creation of learning habits, cognitive and non-cognitive.

Keywords: non-cognitive skills, learning diagnostic, personalization

Workshop: Building the Learning Analytics Curriculum

Charles Lang
Teachers College, Columbia University
charles.lang@tc.columbia.edu

Stephanie D. Teasley
School of Information, University of Michigan
steasley@umich.edu

John Stamper
Human-Computer Interaction Institute, Carnegie Mellon University
jstamper@cs.cmu.edu

ABSTRACT: Learning Analytics courses and degree programs both on- and offline have begun to proliferate over the last four years. Building on the success of the LAK17 curriculum workshop, we plan to work through useful instructional practices, including innovations in classes and programming, similarities and differences for residential vs. online curriculum and a standard learning analytics classroom activity that can be implemented across institutions. We again aim to foster synergy between administrators, instructors, and researchers which can benefit the field as a whole.

Keywords: curriculum, learning analytics instruction, curriculum development, teaching

1 BACKGROUND

Learning Analytics courses and degree programs both on- and offline have begun to proliferate over the last four years. Tens of thousands of students registered for online courses such as DALMOOC (Siemens, Rose, Gasevic, & Baker, 2014) and BDEMOOC (Baker, 2015), there has been considerable excitement around the growth of new degree programs such as the MSc in Learning Analytics at Teachers College, Columbia (Teachers College, Columbia University, 2015) and the strong analytics focus of the MSc in Digital Education at the University of Edinburgh (The Moray House School of Education, The University of Edinburgh, 2015), and also the addition of two new learning analytics degree programs, the EdM in Learning Analytics at Northeastern (2017) and the Graduate Certificate in Learning Analytics at Brandeis University (2017). as well as an explosion of individual courses including those at the University of Michigan (McKay, 2016), the University of South Australia (Rogers, 2016) and others. These efforts are building on the longstanding teaching of influential material at Carnegie Mellon University and the lasting success of the Learning Analytics Summer Institutes (Society for Learning Analytics Research, 2016) as well as professionally- minded projects such as LACE (LACE Project, 2016). As such, the growth of the field is clearly happening not just through research, but in the classroom as well.

1.1 Building on Last Year's Event

The LAK17 workshop on building the learning analytics curriculum provided a space for instructors, administrators and researchers to discuss and debate the many and varied goals of teaching in learning analytics. Speakers from across the US, Canada, Europe and Asia discussed high-level topics such as program creation, ethical considerations and course selection and details such as appropriate assessment, data resources and topic selection. From this discussion we have selected four topics to build on for LAK18 that span these different levels of analysis: 1.) **Similarities & differences between online (MOOC) and residential/in-person curricula**, 2.) **Issues in designing student selection and development models for learning analytics programs,** 3.) **Data resources**, and 4.) **Development of a standard learning analytics classroom activity for introductory classes.**

1.2 Relevance to LAK18 Theme

The inclusion of the user in the analytic design process requires both the dissemination of analytic methods to as broad a community as possible and the ability to teach leaning analytics experts to practice user-centered processes. In keeping with the theme of LAK18, we plan to address how teaching LA can meet these goals. In particular, we will address the inclusion of learner-centered design as a topic in curricula and how it can be incorporated in learning analytics classroom activity.

2 ORGANIZATIONAL DETAILS

2.1 Type of Event

Workshop

2.2 Proposed Schedule

Table 1: Proposed workshop schedule.

Time	Topic
20min	Introduction and review of LAK17 workshop
30min	Student selection and development models
30min	Similarities & differences between online and residential/in-person curricula
20min	Break
30min	Standard learning analytics classroom activity
20min	Data resources
30min	Review and planning for next year

3 OBJECTIVES & OUTCOMES

The overall objectives of the workshop are to discuss practice and share resources for the formal teaching of learning analytics. This will be broken down into four key sections:

3.1 Student selection and development models in established programs

The goals of Learning Analytics programs and courses are many and varied. From providing an introduction to the field for the general public to training people for particular roles within the burgeoning employment market for educational data scientists (or whatever these individuals come to be called). The objective of this section of the workshop will be to consider and discuss the underlying purpose of a learning analytics curriculum at both the program and course level, with particular reference to the types of students these programs are selecting and how they seek to develop their students. This section of the workshop will focus on discussing how course offerings influence the field overall, and how to balance the needs of different stakeholders such as employers, university administrators and researchers in defining the expertise that graduates should hold.

3.2 Similarities & differences between online and residential/in-person curricula

The second objective of the workshop will be to discuss and develop ideas around content and the sequence in which that content should be taught. In particular we hope to discuss the differences between online and in-person content. We also hope to discuss how the teaching of user-centered design differs between these formats. We will work with participants to determine the level of consensus that exists within the community about what content should be included within Learning Analytics programming and what the priorities are across different cultural, geographic, research paradigms and educational settings might be.

3.3 Sharing Tools & Resources

The third aim of the workshop will be to demonstrate useful tools and resources for teaching Learning Analytics. We plan to have an introduction to the use of data resources, out of the box analyses tools and handbooks, and engage in discussion of the pros- and cons- of various programming languages such as Python and R.

3.4 Standard learning analytics classroom activity

Possibly the most challenging part of teaching Learning Analytics is “walking the talk“, or in other words, thinking through how to incorporate Learning Analytics into the courses we teach and programs we run. In many cases, the courses taught by researchers become the first attempts at observing learning analytics in the wild but whether this ad hoc approach can be systematized remains to be seen. Building on the LAK17 workshop we plan to continue our collective work towards a developing a focus activity that would be useful for many courses and contexts to teach Learning Analytics using Learning Analytics.

3.5 Outcomes

We intend outcomes for the workshop to be both community oriented and tangible. The primary outcome will be to engage and support a community that can communicate about problems of practice around the teaching of Learning Analytics. Tangible artifacts of this community will be, as with the LAK17 workshop (http://ceur-ws.org/Vol-1915/paper7.pdf), invited speakers will publish written versions of their presentations in the CEUR Proceedings database. We will also make available our list of useful data sources and the protocol for the standard learning analytics classroom activity.

4 REFERENCES

Baker, R. (2015, May 22). Big Data in Education. Retrieved October 13, 2016, from https://www.edx.org/course/big-data-education-teacherscollegex-bde1x

LACE Project. (2016). Learning Analytics Community Exchange. Retrieved October 14, 2016, from http://www.laceproject.eu/lace/

McKay, T. (2016, August 26). Practical Learning Analytics. Retrieved October 13, 2016, from https://www.edx.org/course/practical-learning-analytics-michiganx-plax-0

Rogers, T. (2016). Learning Analytics and Digital Learning. Retrieved October 13, 2016, from http://programs.unisa.edu.au/public/pcms/course.aspx?pageid=161021

Siemens, G., Rose, C., Gasevic, D., & Baker, R. (2014, June 2). Data, Analytics and Learning. Retrieved October 13, 2016, from https://www.edx.org/course/data-analytics-learning-utarlingtonx-link5-10x

Society for Learning Analytics Research. (2016). LASI – Learning Analytics Summer Institute. Retrieved October 13, 2016, from http://lasi.solaresearch.org/

Teachers College, Columbia University. (2015). The Masters of Science Degree in Learning Analytics. Retrieved October 13, 2016, from https://www.tc.columbia.edu/human-development/learning-analytics/

The Moray House School of Education, The University of Edinburgh. (2015). Master of Science in Digital Education. Retrieved October 13, 2016, from http://digital.education.ed.ac.uk/

Co-creating Dashboards by Aggregating Concept Maps

Ville Kivimäki
Aalto University
ville.kivimaki@aalto.fi

Joonas Pesonen
Rapida Ltd.
joonas.pesonen@rapida.fi

ABSTRACT: Learning analytics dashboards have been used to make students more aware of their learning process. However, being aware does not necessarily lead to a change in behavior. In addition, other tools should complement dashboards. We introduce a method, where curriculum-based concept maps are used for students' self-assessment. Data of these concept maps is then extracted, aggregated and visualized as dashboards, thus creating a feedback loop between curriculum design and students' self-assessment. During the workshop, we demonstrate our method by co-creating a learning analytics dashboard. Concept maps created and updated by workshop attendees will be used as a data source.

Keywords: dashboard, concept map, co-creation, self-assessment, curriculum design, knowledge building, learning analytics

1 WORKSHOP BACKGROUND

Analytics dashboards, where data about learning is presented in a visual form, are a central element in learning analytics. It is suggested that dashboards can enhance self-regulation of learning process and design of learning experiences, yet more research evidence in real-life settings is needed (Klerkx, Verbert & Duval 2013). In a meta-analysis of research on learning analytics dashboards, Jivet, Scheffel, Drachsler and Specht (2017) noticed that most dashboard designs make learners aware of their learning process but fail to use this awareness to improve cognitive, behavioral or emotional competencies. They suggest that "different tools should complement dashboards and be seamlessly integrated in the learning environment and the instructional design" (Jivet et al. 2017).

One widely used tool category to support learning process and evaluation of learning are concept maps (Novak 2008), mind maps (Buzan 1996) and other node-link type knowledge mapping tools. These different formats can be used in complementary ways to enhance motivation, attention, understanding and recall (Eppler 2006). They can also be used to visualize curriculum content and structure (Willcox & Huang 2017). Using methods of social network analysis and graph theory, data can be quantified (McLinden 2013) and then used in learning analytics activities.

In this workshop we introduce a method, where a learning analytics dashboard is co-created by teachers and students using concept maps. Each course's learning content is described with 5-15 key concepts picked by the teacher of the course. These concepts, arranged according to the curriculum structure, form a concept map template that is given to students. During the course of their studies, students self-assess their knowledge level and emotional reactions towards these concepts with icons in the concept map. Students are also encouraged to find cross-links, i.e. relations, between concepts in different courses and name those connections.

Concept map data is regularly quantified to allow further analysis. Aggregated data from multiple concept maps allows us to build a learning analytics dashboard that focuses on knowledge mastery and allows students, teachers and administration to gain insight about how knowledge building happens on course, degree and institution level. The data can then be used in instructional design and curriculum design.

Our work has started from a need to create a dynamic feedback system that would result in weekly feedback rounds between students and teachers. Further development of this method, technology and analysis leads towards more adaptive learning solutions, linking this theme to Grand Challenges for Engineering, i.e. to Advance Personalized Learning (Chase 2008).

2 ORGANISATIONAL DETAILS OF PROPOSED EVENT

2.1 Type of event

Workshop.

2.2 Proposed schedule and duration

Half-day.

2.3 Type of participation

Open.

2.4 Workshop activities

Co-creation of a learning analytics dashboard before and during the workshop.

2.5 The workshop/tutorial activities that participants should expect

2.5.1 Before the workshop

Workshop team (i.e. authors) will produce a concept map template based on LAK18 conference themes and programme. Participants are asked to assess their emotional reaction and knowledge level on these topics. We also ask participants to find cross-links between different parts of the concept map.

2.5.2 During the workshop

Introduction: Participants will be asked to go next to those who have returned their concept map before. (15 minutes)

Group work 1: Small groups update the concept maps by adding their knowledge level and emotional reactions and finding cross-links between concepts. (30 minutes)

Presentation: Workshop team gives a presentation on the overall process of co-creating dashboards using concept maps. Perspectives of student, teacher, programme leader and administration are considered. (30 minutes)

Group work 2: Small groups ideate new prototypes utilizing presented tools and methods. Working method is based on Stanford's Design Thinking Process (Plattner 2010). Groups are reformed so that participants choose their point of perspective: student, teacher, programme leader or administration. Groups are asked to brainstorm and come up with novel ideas for prototypes. (1 hour = 15 minutes brainstorming/ideating + 5 minutes iteration round + 5 minutes choosing one prototype + 25 minutes building prototype concept + 10 minutes preparing for presentation)

Group presentations: (5 minutes for presentation + 5 minutes for discussion) x 4 = 40 minutes

Final remarks and discussions (5 minutes)

2.5.3 After the workshop

Participants will receive dashboard files based on the data they provided before and during the workshop.

2.6 Expected participant numbers and planned dissemination activities to recruit attendees

We expect to have 12 - 24 participants. We plan to disseminate information about the workshop in social media. We aim to engage potential participants through communities around themes like learning analytics, concept maps, mind maps and dashboards.

2.7 Required equipment for the workshop

We require a lecture room with enough space to accommodate 24 participants. It would be best if tables and chairs could be rearranged for the workshop and during the workshop. We also need a screen and a projector, preferably HD quality with HDMI connections. We also need flipchart paper, markers and Post-It notes.

3 WORKSHOP OBJECTIVES OR INTENDED OUTCOMES

1. Participants learn to identify the stages needed to build co-created dashboards from concept map kind of data.

2. Participants learn to operate with concept maps from user's point of view.

3. Participants develop new ideas around the theme.

4. Participants design a concrete prototype

5. Participants demonstrate their own prototype

Overall objective is to develop new concepts, tools or usages for the demonstrated concept through co-creation and ideation. As the workshop will use LAK18 conference program as its material, the dashboard and the datasets will be published during the conference. We will be working under the hashtag #conceptmapanalytics and homepage for conceptmapanalytics is http://blogs.aalto.fi/conceptmapanalytics.

REFERENCES

Buzan, T., & Buzan, B. (1996). The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential.

Chase, V. (edit.) (2008). Grand Challengers for Engineering. (pp. 45-47) National Academy of Sciences, on behalf of the National Academy of Engineering. Available: http://www.engineeringchallenges.org/File.aspx?id=11574&v=34765dff

Eppler, M. J. (2006). A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. Information visualization, 5(3), 202-210.

Plattner, H. (2010). An Introduction to Design Thinking. Available: https://dschool-old.stanford.edu/sandbox/groups/designresources/wiki/36873/attachments/74b3d/ModeGuideBOOTCAMP2010L.pdf

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: pitfalls of learning analytics dashboards in the educational practice. In European Conference on Technology Enhanced Learning (pp. 82-96). Springer, Cham.

Klerkx, J., Verbert, K., & Duval, E. (2014). Enhancing learning with visualization techniques. In Handbook of research on educational communications and technology (pp. 791-807). Springer New York.

McLinden, D. (2013). Concept maps as network data: analysis of a concept map using the methods of social network analysis. Evaluation and program planning, 36(1), 40-48.

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them.

Willcox, K. E., & Huang, L. (2017). Network models for mapping educational data. Available: http://kiwi.mit.edu/papers/education-mapping-network-model-willcox-huang.pdf

Workshop on Non-Cognitive Assessments at Scale for Online Learning

Yuan Wang
Arizona State University
elle.wang@asu.edu

Maria Ofelia Z. San Pedro
ACT, Inc.
sweet.san.pedro@act.org

Srecko Joksimovic
University of South Australia
srecko.joksimovic@unisa.edu.au

Tamera McKinniss & Jason Way
ACT, Inc.
tamera.mckinniss@act.org, jason.way@act.org

ABSTRACT: The importance of fostering and measuring non-cognitive skills, commonly viewed as critical personal attributes necessary for success in classroom, labor market, and life in general, has been widely recognized. However, despite technological advances and emergence of learning analytics as a multidisciplinary research field, the development of non-cognitive skills assessments remains rather limited. Existing research and practice of online learning, focuses primarily on advancing and automating methods of cognitive skills assessments, whereas non-cognitive skills assessments, on the other hand, traditionally rely on self-reported surveys with limited response rates and questionable validity and reliability. This workshop aims at bridging the scarcity of scalable non-cognitive skills assessments in digital educational environments for learning at scale. We start by reviewing different approaches to the non-cognitive skills assessment in various educational contexts, ranging from traditional face-to-face to online learning settings. Further, by leveraging theoretical and technological advances, we discuss the potential for developing innovative methods that would enable assessment of non-cognitive skills in online learning environments with particular focus on learning at scale.

Keywords: Learning Analytics, non-cognitive assessments, online learning

1 WORKSHOP BACKGROUND

1.1 Significance of Non-Cognitive Skills

Educational research and practice show that success within school, workplace, and other aspects of life in general, depend on both cognitive and non-cognitive skills (Camara, O'Connor, Mattern, Hanson, M. A., 2015; Duckworth & Yeager, 2015; Heckman, Stixrud & Urzua, 2006). Although rather challenging to define, non-cognitive skills (Heckman & Rubinstein, 2001) refer to social and emotional learning capabilities beyond academic knowledge. Non-cognitive skills consist of a diverse range of aspects including conscientiousness (Camara et al., 2015), growth mindset (Dweck, 2006), time management, and self-regulation (Gutman & Schoon, 2013), to name a few.

It has been shown that the predictive power of non-cognitive skills on academic achievement across wide range of educational settings is at least equal to or better than the predictive power of cognitive skills. Specifically, non-cognitive skills have been associated with various academic outcomes (McAbee, Oswald, & Connelly, 2014;

Farkas, 2003; Lleras, 2008; Duckworth & Seligman, 2005) including persistence in postsecondary settings with attendance and retention (e.g., Credé, Roch, & Kieszezynka, 2010) and engagement (McClenney, Marti, & Adkins, 2006). Moreover, non-cognitive skills have also been linked to career advancement (Bowles & Gintls, 2002; Heckman et al., 2006), well-being (Cohen, 2006; Strickhouser, Zell, & Krizan, 2017) and key 21st century competencies, such as critical thinking and problem solving (Buckingham Shum & Crick, 2016).

1.2 Challenges of Non-Cognitive Assessments *at Scale* Online

The availability of clickstream data from large-scale online learning platforms, such as Massive Open Online Courses (MOOCs), are expected to enable assessments at scale to advance the science of learning and learner success (Stokes, 2013). Although assessment of learner success should include both cognitive and non-cognitive aspects, more attention has been given toward measuring learners' cognitive abilities, oftentimes operationalized by test scores or the number of clicks in the environment (Reich, 2015), than on the non-cognitive abilities.

Much of the research studies on non-cognitive assessments are focused on theory development rather than actionable applications (Duckworth & Yeager, 2015) derived from practical measurements incorporating short-term feedback on progress toward improvement (Bryk et al., 2015). The lack of feedback in assessment limits applicability. A recent review of MOOC literature (Joksimovic et al., 2017) shows that existing research focuses primarily on understanding factors that explain academic learning outcomes, thus in most cases failing to include the assessment of non-cognitive skills. In the following section, we identify recent development of online non-cognitive assessments opportunities toward actionable applications.

1.3 Recent Development of Online Non-Cognitive Assessments

Survey Measures. Researchers (e.g., Kizilcec and Schneider, 2015) investigated the relationship between non-cognitive factors, such as learner motivation and self-efficacy, and learning outcomes using self-reported survey items. They found that different motivational types lead to varied types of engagement patterns. While these studies provide critical insights toward theory development, applying self-reported surveys alone limits what we can learn about online learners' non-cognitive capabilities. For example, self-reported surveys often have limited response rates, lower than 10 percent in MOOCs (e.g., Wang & Baker, 2015). Moreover, it is also arguable to what extent self-reported data provide unbiased data (Zhou & Winne, 2012). Consequently, not reliably assessing online learners' non-cognitive skills *at scal*e undermines the extent to which learning analytics can be applied to assess and improve student learning.

Learning Analytics (LA) and Educational Data Mining (EDM). In addition to traditional survey measures, recent advances in the LA and EDM research fields provide innovative venues for investigation of students' non-cognitive skills (e.g., emotion, metacognition, motivation). Multimodal data, such as eye gazes, facial expressions of emotions, heart rate and electro-dermal activities (D'Mello, Dieterle, & Duckworth, 2017) were investigated. In addition, toward scaling up, EDM researchers have developed "automated detectors" out of computer log data to infer student's emotions and engagement in real-time. These detectors, developed from a combination of expert field observation (e.g. Ocumpaugh, Baker, & Rodrigo, 2012) and data mining on log files, can accurately predict affect and engagement (Baker .et al, 2012).

LA and EDM methods offer promising paths toward assessing online learners' non-cognitive skills, and ultimately toward deriving actionable applications to improve learner non-cognitive skills and subsequently improve cognitive outcomes as well. The direction toward developing tools and metrics to assess online learners' non-cognitive skills reliably and at scale remain an ongoing focus.

2 RELEVANCE TO LEARNING ANALYTCIS & KNOWLEDGE (LAK)

In line with LAK18's theme of *human factors* in LAK systems, the proposed workshop focuses on assessment of non-cognitive skills in order to understand the *tools* and *metrics* that allow us to move beyond commonly used approaches. Specifically, we aim at uncovering theory-informed data-driven aspects of online learning settings in general and learning at scale in particular. Therefore, this workshop is relevant to the learning analytics community in the following ways:

1) *Enriching assessments through learning analytics*: Leading learning analytics researchers raised the concern that non-cognitive measures are critical and need to be incorporated into the design of learning analytics tools in addition to cognitive measures (e.g., Lonn, Aguilar & Teasley, 2015). Building on the advances in multimodal learning analytics, the proposed workshop will address this through exploration of innovative ways to measure non-cognitive skills at scale.
2) *Connecting learning analytics with authentic contexts* – Learning analytics should be in service of building 21st century competencies which include multiple key non-cognitive skills (Buckingham Shum & Crick, 2016) that help learners to grow beyond academic settings. The direction of non-cognitive assessments *at scale*, via incorporating LA, taps right into the connection between short-term learning goals and long-term life goals including career development, well-being, and life-long learning abilities.

3 OBJECTIVES

The theme of the proposed workshop underscores challenges and needs of assessment of non-cognitive capabilities at scale in online learning environments while aiming to deriving actionable insights. Presentations on finished, ongoing, and proposed studies, as well as facilitated discussion sessions are planned to develop a preliminary framework to illustrate the current development of non-cognitive assessments within online learning environments and to inspire attendees to come up with exciting new lines of research of their own. Toward building this framework, the following guiding questions are included:

- **Theories:** What are the relevant theories that can inform assessments of non-cognitive skills and their application?
- **Data sources:** Where and how can data related to learners' non-cognitive skills be measured and collected?
- **Tools:** What analytical and assessment tools are useful in analyzing non-cognitive constructs at scale? How can we appropriately link non-cognitive assessment with cognitive assessment?
- **Methods:** What analytical methods have been used? What other methods can be applied? How are non-cognitive constructs operationalized?
- **Generalizability:** What kind of practices and findings are domain-general across learning environments?
- **Applicability:** How can research findings translate into actionable insights for various stakeholders (learners, instructors, administrators, investors, etc.)?

4 ORGANIZATIONAL PLANS

Type of event: Workshop
Proposed schedule and duration: Half-day

Type of participation: Mixed participation. The proposed workshop plans to send out a call for submissions to select the best peer-reviewed paper. To ensure the quality of the submission, we have confirmed with a group of leading experts in the fields of learning analytics, learning sciences and psychology to serve as our program committee members if accepted.

Proposed activities: Keynote presentations from industry and academia; peer-reviewed presentations; guided small-group discussions; Birds of a feather

Expected participant numbers: 35-40

Planned dissemination activities to recruit attendees

- **Workshop Proceedings:** We plan to publish in workshop proceedings with a free open-access publication service such as CEUR-WS.org.
- **Social Media:** We will create a workshop website and use #noncogLAK18 as the primary hashtag to encourage discussions and communication through social media channels such as yellowdig.com.

- **Website:** We also plan to integrate all relevant resources, produced before and during the workshop such as presentation slides and discussion notes, on the website to encourage ongoing communication and collaboration after the workshop.
- **Required equipment for the workshop:** A conference room with capacity for up to 40 people with a setup that allows for small group discussions. A computer and a screen for presentations are also needed.

REFERENCES

Baker, R.S.J.d., Gowda, S.M., Wixon, M., Kalka, J., Wagner, A.Z., Salvi, A., Aleven, V., Kusbit, G., Ocumpaugh, J., Rossi, L. (2012) Towards Sensor-free Affect Detection in Cognitive Tutor Algebra. *Proceedings of the 5th International Conference on Educational Data Mining*, 126-133

Barrick, M. R. , & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1-26.

Buckingham Shum, S., & Crick, R. D. (2016). Learning Analytics for 21st Century Competencies. *Journal of Learning Analytics*, *3*(2), 6-21.

Bowles, S., & Gintis, H. (2002). Schooling in capitalist America revisited. Sociology of Education, 75(1), 1-18.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to Improve: How America? s Schools Can Get Better at Getting Better*. Harvard Education Press. 8 Story Street First Floor, Cambridge, MA 02138.

Camara, W., O'Connor, R, Mattern, K, Hanson, M. A. (2015). Beyond Academics: A Holistic Framework for Enhancing Education and Workplace Success. (ACT Research Report 2015-4). Retrieved from http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-4.pdf

Cohen, J. (2006). Social, emotional, ethical, and academic education: Creating a climate for learning, participation in democracy, and well-being. *harvard educational Review*, *76*(2), 201-237.

Credé, M., Roch, S. G., & Kieszczynka, U. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *80*, 272–295.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist*, *52*(2), 104-123.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological science*, *16*(12), 939-944.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, *92*(6), 1087.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237-251.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House Incorporated.

Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. Annual Review of Sociology, 29, 541-562. doi: 10.1146/annurev.soc.29.010202.100023

Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. *Education Empowerment Foundations, London*. Retrieved from: http://hdl.voced.edu.au/10707/287500

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*, *91*(2), 145-149.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411-482.

Joksimović, S., Poquet, S., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A. C., & Brooks, C. (2017). How do we model learning at scale? *A systematic review of the literature. Review of Educational Research*.

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. Social Science Research, 37(3), 888-902.

Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, *47*, 90-97.

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *22*(2), 6.

Lee, K., & Ashton, M. C. (2004). Psychometric prosperities of the HEXACO personality inventory. *Multivariate Behavioral Research*, *39*, 329–358.

McAbee, S. T., Oswald, F. L., & Connelly, B. S. (2014). Bifactor models of personality and college student performance: A broad vs. narrow view. *European Journal of Personality*, *28*, 604-619. doi:10.1002/per.1975

McClenney, K., Marti, C. N., & Adkins, C. (2006). *Student engagement and student outcomes: Key findings from CCSSE validation research*. Austin, TX: Center for Community College Student Engagement.

Ocumpaugh, J., Baker, R.S.J.d., & Rodrigo, M.M.T. (2012). *Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0.*Technical Report. New York, NY: EdLab.

Reich, J. (2015). Rebooting MOOC research. *Science*, *347*(6217), 34-35.

Robbins, S. B., Allen, J., Casillas, A., Peterson, C., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology*, *98*, 598–616.

Stokes, P. (2013). The Particle Accelerator of Learning. *Inside Higher Ed*, 22.

Wang, Y., & Baker, R. (2015). Content or platform: Why do students complete MOOCs?. *Journal of Online Learning and Teaching*, *11*(1), 17.

Strickhouser, J.E., Zell,E., & Krizan, Z. (2017). Does personality predict health and well-being? A metasynthesis. *Health Psychology,* 36(8), 797-810.

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*(6), 413-419.

Rethinking Non-Cognitive

Thomas G. Fikes
EdPlus Action Lab, Arizona State University, Scottsdale AZ 85257, USA
thomas.fikes@asu.edu

1 INTRODUCTION

The way we talk about things matters. This fact is central to what we could call the non-cognitive revolution in education, particularly when it comes to self-talk: academic mindset (a quintessential non-cognitive factor) involves how students talk to themselves about their abilities, failings, and ability to change. Of course, the talking doesn't have to be aloud – "talking" here stands proxy for thinking, or at least a particular kind of thinking, in which one *explicitly* considers (internally or aloud) one's *beliefs*. Metacognition – "thinking about one's own thinking"— is another example of self-talk that matters. And when learners engage in the self-talk that undergirds mindset and metacognition in the presence of others, that talk becomes even more powerful as social processes amplify the effects: "math is hard" and "I'm not a math person", for example, become normative within a group, forming a kind of self-fulfilling prophecy. The point here is that how we talk matters, and that it can matter even more profoundly in a social context.

In this short essay I explore some of the consequences of talking about "non-cognitive" as a community of researchers, offer a solution to what I see as a problem inherent in the use of that term, and introduce the seven papers that comprise the LAK 2018 *Workshop on Online Learning & Non-Cognitive Assessment at Scale*. Although my comments here are not specific to online learning per se, the problems I highlight (and therefor our solutions to those problems) have a disproportionate impact on research at the intersection of non-cognitive and online learning: because the digital modality is novel, we stand in particular need of conceptual, theoretical, and methodological clarity.

2 CONSEQUENCE 1: THE NON-COGNITIVE REVOLUTION AS DISRUPTER

One important consequence of introducing the concept of "non-cognitive" into the educational research literature was to highlight the poverty of the then standard model of academic learning. In that model, domain-specific knowledge and skills ("content knowledge") was understood to be the result of domain-specific curriculum, learning behaviors, and teaching practices; and in turn, career and work success were mediated by these domain-specific acquisitions. Strangely, "domain specific" and "content knowledge" had come to be known as "cognitive", inferring a kind of scientific credibility following the success of the cognitive revolution. The (re)introduction of non-cognitive (e.g., by Heckman & Rubinstein, 2001) served an important disruptive function of reintroducing to the scientific discourse on education what educators and psychologists already knew: "It is common knowledge outside of academic journals", Heckman and

Rubinstein reminded us, “that motivation, tenacity, trustworthiness, and perseverance are important traits for success in life” (p. 145). Those same academic journals are now filled with studies of these factors and many others like them.

Education research was not alone in this respect: many other fields in the brain, behavioral, and social sciences have been reclaiming ground lost in an overzealous use of “cognition” following the cognitive revolution of the 1970s. In a paper that redirected the course of the scientific study of emotions, for example, Izard (1993) reminded us that the activation of emotions entails not only cognitive processes (such as appraisal), but also non-cognitive processes; and that “information processing” was largely non-cognitive by any reasonable understanding of the term “cognition.”

Unlike Izard’s careful unpacking of cognitive and non-cognitive in the emotions literature, however, the field of educational research has used “non-cognitive” in ways that obfuscate important theoretical concepts, leading to a confused literature on the subject that impedes theoretical, empirical, and practical advances. This second consequence partially undermines the disruptive advantaged gained in the first.

There is confusion caused by the continued use of “non-cognitive” to describe the collection of non-domain-specific factors and processes that contribute to academic performance and the career, work, and life success that can result from academic success, and the implicit association of domain-specific knowledge and skills with “cognitive”.

3 CONSEQUENCE 2: THE NON-COGNITIVE REVOLUTION AS OBFUSCATOR

The problem with our use of “non-cognitive” in the educational literature to refer to the diversity of factors and skills under consideration is twofold: first, these factors and skills are inherently cognitive; and second, grouping them into a single category leads to theoretical incoherence. Such confusions are misleading and counterproductive.

3.1 The Inherently Cognitive Nature of Non-cognitive Factors

Consider problem solving and metacognition. These are obviously deeply cognitive concepts, forming central fields of inquiry across the history of cognitive psychology, cognitive science, and related academic fields. Similarly, academic behaviors, academic perseverance, social skills, and learning strategies are habits, actions, and outcomes that fundamentally rest on cognition as well as basic “non-cognitive” motivational and affective processes, reinforcement histories, and sociocultural contexts. These “non-cognitive” functions are deeply and primarily cognitive in nature.

Academic mindset, self-efficacy, sense of belonging to an academic community, beliefs associated with expectancy-value theory form a separate group of so-called “non-cognitive” factors. What separates them from the first group is their inherently affective (or motivational) nature, and so perhaps these have more claim to the non-cognitive label – and yet, they are inherently cognitive as well, resting as they do on beliefs, expectancies, and attentional and memory processes. These might fruitfully be thought of as attitudes, typically defined by social psychologists as comprising cognitive, affective, and behavioral components (e.g., Eagly & Chaiken, 1998), or as sentiments ("functional networks of attitudes and emotions"; Gervais & Fessler, 2017, p. 1).

3.2 Theoretical Incoherence

By separating out factors that contribute to academic success into theoretically coherent categories, we can more clearly see similarities and differences. Such theorizing has practical consequences, since interventions that are successful with respect to one factor may also enjoy success for factors that share causal and functional mechanisms, but be time-wasters for dissimilar factors. Theoretical clarity can also point out whole regions of missing territory in the theoretical landscape; sociocultural context (e.g., Markus & Kitayama, 1991), for example, may represent a very fruitful direction in academic success research, highlighting important factors beyond those described in today's "non-cognitive" taxonomies. This is particularly important for the digital education world, which frequently crosses cultural borders as it serves a global clientele. Finally, a theoretically coherent framework can aid in moving academic success research into new territory: in our lab (https://edplus.asu.edu/projects/action-lab), which focuses on digital teaching and learning and online educational programs in higher education, insight into the causal processes that underlie academic success would aid in the discovery of innovations in pedagogy, policy, and technology that allow us to close the gaps in digital education and provide access to quality higher education at scale.

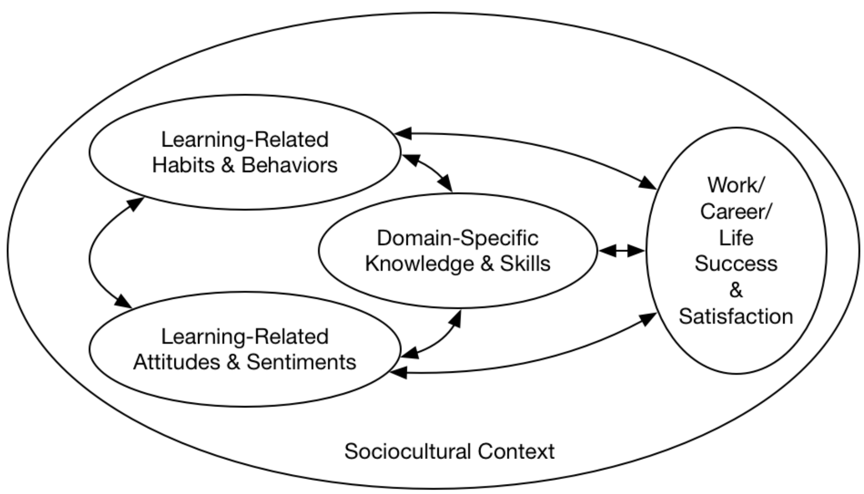


Figure 1: A framework for rethinking non-cognitive skills and factors that contribute to academic success.

4 A RETHINKING

In Figure 1, I present one possible theoretical framework for helping to rethink non-cognitive. I've recast "non-cognitive" factors and skills into three theoretical categories: *learning-related habits* (I use the old Jamesian language of "habits" here, which came back into vogue at in the 1980s, because it explicitly integrates cognition and behavior), *learning related attitudes*, and *sociocultural context*. Learning-related habits and attitudes interact dynamically with domain-specific knowledge and skills (cf. content knowledge) and with "ultimate outcomes" for which academic experience is supposed to prepare us (such as success and satisfaction in work and career). Consider, in contrast, the pre-non-cognitive-revolution model of learning: domain-specific learning habits (and curriculum and teaching) drives domain-specific knowledge and skills, which in turn drives later success. Those domain-specific habits and the knowledge and skills they support are what (somehow) came to be known as "cognitive". Such domain-specific habits would live in the category of Learning-Related Habits and Behaviors in Figure 1. The present framework adds in a host of other "non-cognitive" habits (problem solving, metacognition, social skills, etc.), "non-cognitive" attitudes (mindset, self-efficacy, etc.), and sociocultural context (e.g., individualist vs. collectivist). Such a framework would not replace extant models or frameworks (see the Farrington et al., 2012, collection for one very useful example). Rather, it provides (a) a means for us to incorporate extant work in a way that honors the very "cognitive" nature of the "non-cognitive" literature in education, (b) a theoretically coherent way of talking about the factors that contribute to academic success, and thus (c) a means of practically working with such factors and skills to innovate education at scale.

Table 1: LAK 2018 Workshop on Online Learning & Non-Cognitive Assessment at Scale papers and "non-cognitive" factors considered.

Workshop Paper	Habits, Behaviors, and Skills	Attitudes and Sentiments
Modell	social skills	
McKinniss, Ofelia Z. San Pedro, Dixon, & Way	persistence, collaboration, problem solving	
Wang, Cunningham, Arcuria, Fikes, & Pugliese	time/project management, leadership, decision-making	
Dillahunt & Wang	acquisition of human capital, personal adaptability	career identity
Turkay, Seaton, & Ang	self-regulation	self-efficacy
Li & Krasny	intercultural competence	civic engagement
McCarthy, Likens, Kopp, Perret, Watanabe, & McNamara	grit, learning orientation, performance orientation	grit?

In Table 1, I list the papers contributing to the workshop, the primary "non-cognitive" factors and skills that are represented in those papers, and a tentative mapping of those onto the habits and attitudes taxonomy in Figure 1 (I omit sociocultural context, since it did not appear that any of the papers addressed factors that fit that category).

Those who pioneered mindset, metacognition, self-efficacy and similar concepts in educational research were right: the way we talk about things matters. I hope that those of us who study these processes in our students can practice what we preach, rethink non-cognitive, and clean up the way we talk about the revolution to which we are contributing.

REFERENCES

Eagly, A. H., & Chaiken, S. (1993). The psychology of attitudes. Orlando, FL: Harcourt Brace Jovanovich College.

Farrington, C.A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T.S., Johnson, D.W., & Beechum, N.O. (2012). Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review. Chicago: University of Chicago Consortium on Chicago School Research.

Gervais, Matthew M. & Fessler, Daniel M. T. (2017). On the deep structure of social affect: Attitudes, emotions, sentiments, and the case of “contempt”. Behavioral and Brain Sciences, 40.

Heckman, J., & Rubinstein, Y. (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program. The American Economic Review, 91(2), 145-149. Retrieved from http://www.jstor.org/stable/2677749

Izard, C. E.. (1993). Four systems for emotion activation: cognitive and noncognitive processes. Psychological review, 100, 68-90. 10.1037/0033-295X.100.1.68.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98(2), 224-253. 10.1037/0033-295X.98.2.224

McCarthy, K., Likens, A., Kopp, K., Perret, C., Watanabe, M., and McNamara, D. (2018). The “LO”-Down on Grit: Non-Cognitive Trait Assessments Fail to Predict Learning Gains in iSTART and W-Pal. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

Wang, E. Y., Cunningham, J., Arcuria, P., Fikes, T.G., and Pugliese, L. (2018). Exploring Non-Cognitive Reasons behind Success after Failure. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

Turkay, S., Seaton, D. and Ang, A. (2018). Itero: A New Tool for Exploring Writing Self-efficacy through Revision History Analytics. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

Modell, M. G. (2018). Evaluating group dynamics and individuals’ contributions to them. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

Li, Y. and Krasny, M. (2018). Enhancing global engagement through environmental education MOOCs. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

McKinniss, T., Ofelia Z. San Pedro, M., Dixon, A. and Way, J. (2018). Non-cognitive Skills: Relevance & Measurement during Online Learning. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

Dillahunt, T., and Wang, E.Y. (2018). Non-cognitive Assessments at Scale: MOOCs and Employability. Paper presented at LAK 2018, Workshop on Online Learning & Non-Cognitive Assessment at Scale, Sydney, NSW, AU.

Evaluating group dynamics and individuals' contributions to them

Micah Gideon Modell
SUNY Korea
micah.modell@SUNYKorea.ac.kr

ABSTRACT: While there's plenty of material on how to design optimal team tasks and high-performing teams, there is little on identifying disruptive group dynamics. If we aren't able to consistently recognize and classify a problem when we see it, how can we possibly help team members to navigate these disruptions, much less turn them to the advantage of the entire team? Conflicts within a team can be strengths (K. A. Smith, Johnson, & Johnson, 1981) especially as participants develop the skills to negotiate resolution and better understand the problems in the process. However, if allowed to fester, they are toxic. They detract from the experience and the product and effectively teach that collaboration is something to be avoided (Modell, 2017).

The CoLab.online project collects regular self- and peer-assessment from collaborative learning group participants. This work aims to understand this data and enable algorithmic identification of patterns that signal disruptive dynamics. The long-term goal of this research is to arm instructors to help students learn to effectively navigate such situations and implement effective resolution behaviors.

Keywords: collaborative learning, learning analytics, self and peer assessment

1 INTRODUCTION

Collaborative group work is a popular student-centered instructional method (Baldwin, Bedell, & Johnson, 1997; Cooper & Robinson, 1998; Gibbs, 1995; Johnson & Johnson, 1998; Paulus, Kohn, & Dzindolet, 2011). Learning to work in teams in the classroom helps to establish support networks (Baldwin et al., 1997), and encourages learning by teaching and discussing the topic with peers (M. K. Smith et al., 2009; Vygotsky, 1978). Collaborative learning prepares students for an increasingly collaborative workplace (Baldwin et al., 1997; Ohland et al., 2012).

In fairness, instructors also use collaborative group projects because they perceive a reduction in grading load as a group of four students will only produce one (often more complex) project (Goldfinch & Raeside, 1990; Tucker & Reynolds, 2006; Young & Henquinet, 2000). This should, in turn, benefit the students by allowing instructors to provide students feedback that is more thorough.

1.1 The problem

However, the ability to contribute constructively to a team is a skill which does not occur naturally and must be nurtured. To develop a workforce that is consistently capable of reaping the benefits of working in teams, it is imperative that our instructors support their groups to help them identify and navigate the problems they encounter These disruptive group dynamics take many forms, including: social loafing (De Vita, 2001; Freeman, 1995; King & Behnke, 2005; Zhang, Johnston, & Kilic, 2008), inappropriate division of labor (Sheingold, Hawkins, & Char, 1984), groupthink (Falchikov, 1995; Haynes, 2012), or group domination (Cohn, Ohlsen, & Proff, 1960; Rogat & Adams-Wiggins, 2014) to name a few (see Modell, 2015 for more).

When disruptive group dynamics emerge, they can serve not only to create discomfort for students, but can actually hinder the learning process (Webb, 1995). As an example, a free riding (or social loafing) student is not completing the practice assigned by their instructor (presumably to hone develop their skills). Unfortunately, the group product does not reflect the capabilities of the individual, but the instructor is generally unaware, and the students share a grade, thereby undermining the credibility of the evaluation.

Unfortunately, research suggests that instructors in higher education struggle to successfully diagnose such disruptive behaviors in groups (Modell, 2015). This is the case for a variety of reasons, including: complexity of group dynamics, inability to comprehensively monitor group interactions, group dynamics being considered peripheral to the content of the course, and a lack of training in managing groups. Further analysis indicates that instructors do not share a professional vision or common conception of how a group should operate or how its health ought to be evaluated (Modell, 2017).

While teamwork is increasingly a fact of post-academic life, our efforts to prepare students to reap the benefits of collaboration are not only often unsuccessful, but actually have the effect of teaching them to avoid such situations in the future. While it would be wonderful for all instructors to be thoroughly trained in the use of this method, it is likely that all involved would benefit from a deeper understanding of these dynamics and the development of tools to assist instructors in their efforts.

2 THE METHOD AND THE DATASET

The CoLab.online platform[1] represents the third iteration of the platform originally described by Modell (2013). It offers instructors the ability to administer weekly self- and peer-assessment to collaborative groups to help them visualize group member perceptions of group dynamics. It provides groupwork simulations to facilitate discussion of how groups operate and what disruptive behaviors look like. It also aids with composing desirable groups and gamifying assigned reading activities. Most important, it captures all this data and provides an interface through which instructors can see how their managed groups are interacting while researchers are able to analyze the same data through an anonymizing interface. My efforts are aimed either at improving the data collected by this platform and better understanding it.

[1] http://www.CoLab.online

2.1 The method

Students in groups administered by the system are asked weekly to allocate contributions according to their perceptions (see Figure 1). Allocations are requested across multiple factors along which individuals can contribute towards positive group dynamics. The method holds the group as the unit of analysis (rather than the individual), and therefore each factor represents a 'zero sum game' such that saying one member was responsible for suggesting 50% of the ideas in a wee, that would leave only 50% for the rest of the team members to split. This renders the evaluations less of a popularity contest. The system currently defaults to use the following set of factors (adapted from Goldfinch & Raeside, 1990):

- Organising the group
- Understanding what was required
- Suggesting ideas
- Coming up with something useful
- Performing tasks allocated by the group (on time!)

However, there is an effort underway to distill a set of factors from a comprehensive review of the literature on assessing group dynamics.

Students are reminded by email to complete their assessments and the form is optimized to make data entry itself quick and easy. In practice, the response rates are quite high (sometimes over 85%).

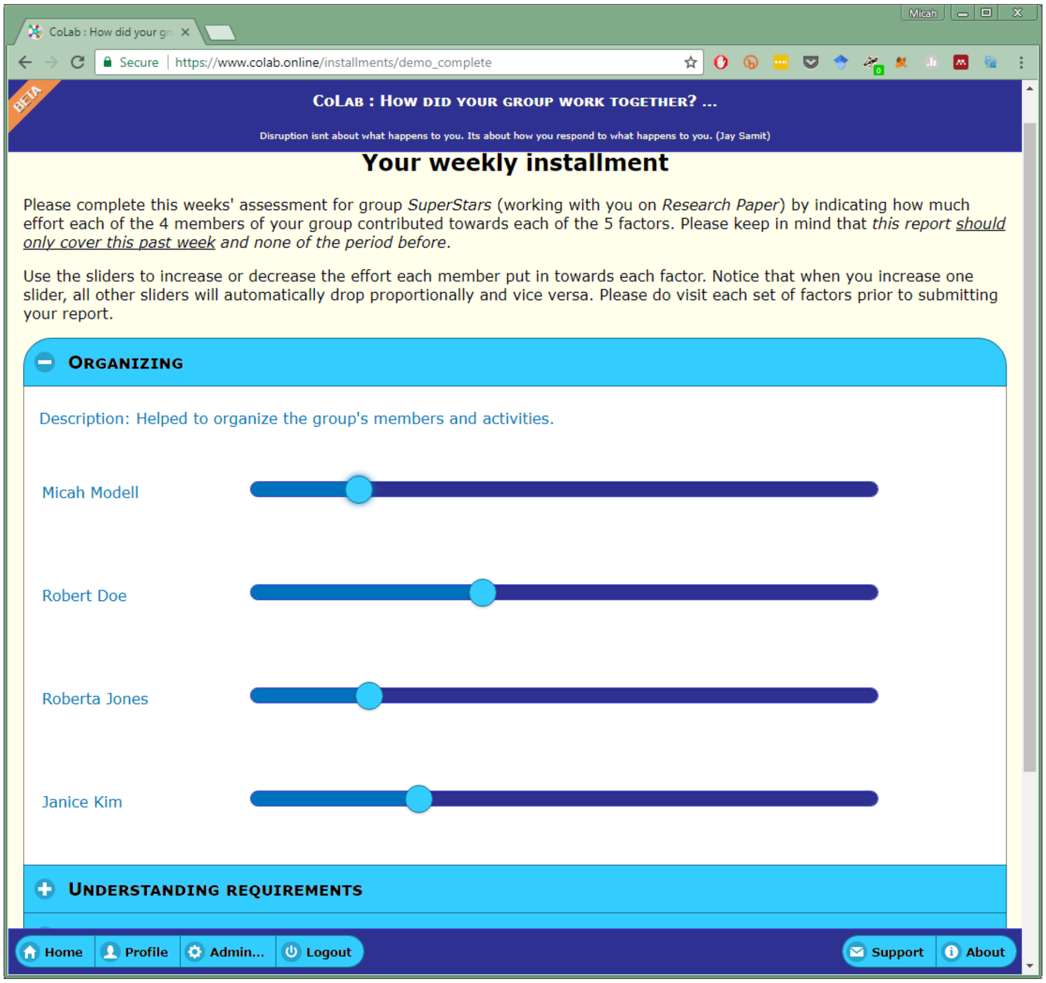


Figure 1: Self- and peer-assessment data entry form.

2.2 The data

Currently, the system exposes the recorded self- and peer-assessment data in the form of a set of charts (see figure 2). At the current time, instructors have the option to view visualizations representing different slices of the data both in aggregate using simple averages. This information enables instructors to see when groups are running into difficulties and even to gain some insight into the nature of those difficulties. Furthermore, this data represents the evolution of the group over time enabling instructors to offer students customized guidance based upon the issues they are experiencing.

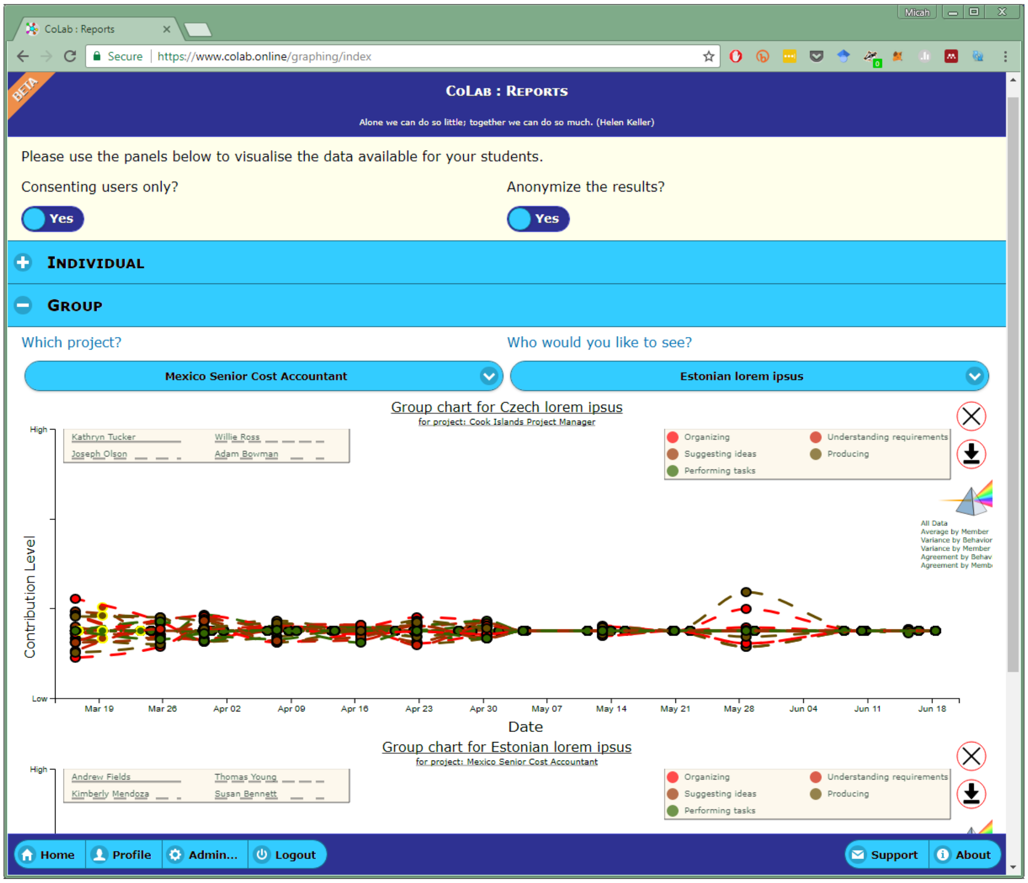


Figure 2: Self- and peer-assessment data reporting.

At this time, data is not made directly available to the students. The primary reason is that doing so would threaten the quality of the data: a measure of anonymity in peer assessment encourages honesty and reflection and yields more accurate results (Thompson & McGregor, 2005; Vanderhoven, Raes, Schellens, & Montrieux, 2012). Furthermore, while the computer excels at the tedious tasks of collecting, processing and visualizing the data, instructors are still required to contextualize it so as to help students learn from the experience and develop coping strategies.

In practice, the data is a bit overwhelming and would benefit from further automated analysis.

3 CURRENT USES AND NEXT STEPS

This system has been live since the spring of 2017 and has been augmented with additional functionality including a group work simulation, a gamified reading activity, and measures of group diversity). It has been used to support both undergraduate and graduate classes engaged in full-semester (14 week) collaborative group projects, offering the instructors using it insight into their students' perceptions of their group's behavior. They (the instructors) are experimenting with the functionality and offering feedback to the researcher in an effort to refine its capabilities and develop best practices. While the captured data would lend itself to the generation of a grade, and instructors have requested this functionality, the researcher promotes use of the data for formative purposes.

This data represents an opportunity to develop a deeper understanding of group dynamics and how individuals contribute to the smooth operation of the groups in which they participate. The system has been developed with research as an integral part of its mission and therefore all personally identifiable information also has an anonymized version. This mean that researchers can review and analyze data using mixed methods at any time including student stories (in comments) without having to be concerned with anonymity. Work is also in progress to make an API accessible to instructors and researchers.

While the visualizations help, the data is still overwhelming, and it is difficult to make sense of it at the moment. Work is currently underway to develop representations of progress over time. Do team members agree on one another's contributions? How does that agreement change over the duration of a course? Most important: what are the indicators of a group having problems and can the system proactively inform instructors of groups experiencing difficulties? Does use of the system yield any change in students' behavior on successive projects? Can we use student self- and peer-assessment data to recommend interventions to help students navigate disruptions? For that matter, this data could also be used to identify situations on which to test experimental interventions.

Research is underway to refine the set of factors assessed by the system. These factors are being derived from a review of the literature on group dynamics and group assessment. It is hoped that the inclusion of a rigorously vetted set of factors will offer deeper insights into group dynamics and a richer language for discussing them.

It may be further possible to use this data, collected longitudinally, and combined with additional measures, to evaluate the development of leadership skills, embrace of diversity, and possibly creative/generative skills. For example, can group participation patterns be used to predict corporate leadership or innovation? Can a perspective-taking intervention smooth out group performance and help students appreciate value from group member diversity?

4 ACKNOWLEDGMENTS

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICT Consilience Creative program (IITP-2017- R0346-16-1007) supervised by the IITP (Institute for Information & communications Technology Promotion).

REFERENCES

Baldwin, T. T., Bedell, M. D., & Johnson, J. L. (1997). The social fabric of a team-based MBA program: Network effects on student satisfaction and performance. *Academy of Management Journal*, *40*(6), 1369–1397.

Cohn, B., Ohlsen, M., & Proff, F. (1960). Roles played by adolescents in an unproductive counseling group.

The Personnel and Guidance Journal, *38*(9), 724–731. https://doi.org/10.1002/j.2164-4918.1960.tb02648.x

Cooper, J., & Robinson, P. (1998). *Small-group instruction in science, mathematics, engineering and technology (SMET) disciplines: A status report and an agenda for the future*. *Journal of College Science Teaching*. Madison, WI.

De Vita, G. (2001). The use of group work in large and diverse business management classes: Some critical issues. *The International Journal of Management Education*, *1*(3), 27–35. https://doi.org/10.3794/ijme.13.11

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International*, *32*(2), 175–87.

Freeman, M. (1995). Peer assessment by groups of group work. *Assessment & Evaluation in Higher Education*, *20*(3), 289–300. https://doi.org/10.1080/0260293950200305

Gibbs, G. (1995). *Assessing student centred courses*. *Assessing student centred courses*. Oxford, UK: Oxonian Rewley Press Ltd.

Goldfinch, J., & Raeside, R. (1990). Development of a peer assessment technique for obtaining individual marks on a group project. *Assessment & Evaluation in Higher Education*, *15*(3), 210–231. https://doi.org/10.1080/0260293900150304

Haynes, N. M. (2012). *Group Dynamics*. Lanham, MD: Univesity Press of America, Inc.

Johnson, D. W., & Johnson, R. T. (1998). Cooperative learning and social interdependence theory. In R. S. Tinsdale, L. Heath, J. Edwards, E. J. Posavac, F. B. Bryant, J. Myers, … E. Henderson-King (Eds.), *Theory and research on small groups* (pp. 9–35). New York: Plenium Press.

King, P. E., & Behnke, R. R. (2005). Problems associated with evaluating student performance in groups. *College Teaching*, *53*(2), 57–61.

Modell, M. G. (2013). Iterating over a method and tool to facilitate equitable assessment of group work. *International Journal of Designs for Learning*, *4*(1), 39–53.

Modell, M. G. (2015). *Distinguishing between healthy and dysfunctional student project teams: An elusive instructor challenge*. Indiana University, Bloomington. Retrieved from http://hdl.handle.net/2022/20351

Modell, M. G. (2017). Instructors' professional vision for collaborative learning groups. *Journal of Applied Research in Higher Education*, *9*(3), 346–362. https://doi.org/10.1108/JARHE-11-2016-0087

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., … Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self- and peer evaluation. *Academy of Management Learning & Education*, *11*(4), 609–630. https://doi.org/10.5465/amle.2010.0177

Paulus, P. B., Kohn, N., & Dzindolet, M. (2011). Creativity Versus Innovation, 446–452. https://doi.org/10.1016/B978-0-12-375038-9.00217-X

Rogat, T. K., & Adams-Wiggins, K. R. (2014). Other-regulation in collaborative groups: implications for regulation quality. *Instructional Science*, *42*(6), 879–904. https://doi.org/10.1007/s11251-014-9322-9

Sheingold, K., Hawkins, J., & Char, C. (1984). "I'm the thinkist, you're the typist": The Interaction of Technology and the Social Life of Classrooms. *Journal of Social Issues*, *40*(3), 49–61. https://doi.org/10.1111/j.1540-4560.1984.tb00191.x

Smith, K. A., Johnson, D. W., & Johnson, R. T. (1981). Can conflict be constructive? Controversy versus Concurrence seeking in learning groups. *Journal of Educational Psychology*, *73*(5), 651–663. https://doi.org/10.1037/0022-0663.73.5.651

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why Peer Discussion Improves Student Performance on In-Class Concept Questions. *Terra Nova*, *323*(5910), 122–124. https://doi.org/10.1126/science.1165919

Thompson, D., & McGregor, I. (2005). Self and peer assessment for group work in large classes. *Architecture*, *30*. Retrieved from http://hdl.handle.net/10453/1373

Tucker, R., & Reynolds, C. (2006). The impact of teaching models, group structures and assessment modes on cooperative learning in the student design studio. *Journal for Education in the Built Environment*, *1*(2), 39–56.

Vanderhoven, E., Raes, A., Schellens, T., & Montrieux, H. (2012). Face-to-Face Peer Assessment in Secondary Education: Does Anonymity Matter? *Procedia - Social and Behavioral Sciences*, *69*(Iceepsy), 1340–1347. https://doi.org/10.1016/j.sbspro.2012.12.071

Vygotsky, L. S. (1978). *Mind in society*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.), *Mind in Society The Development of Higher Psychological Processes*. Harvard University Press. https://doi.org/10.1007/978-3-540-92784-6

Webb, N. M. (1995). Group Collaboration in Assessment: Multiple Objectives, Processes, and Outcomes. *Educational Evaluation and Policy Analysis*, *17*(2), 239–261. https://doi.org/10.3102/01623737017002239

Young, C. B., & Henquinet, J. A. (2000). Viewpoint: A conceptual framework for designing group projects. *The Journal of Education for Business*, *76*(1), 56–60.

Zhang, B., Johnston, L., & Kilic, G. B. (2008). Assessing the reliability of self- and peer rating in student group work. *Assessment & Evaluation in Higher Education*, *33*(3), 329–340. https://doi.org/10.1080/02602930701293181

Non-cognitive Skills: Relevance & Measurement during Online Learning

Tamera L. McKinniss
ACT, Inc.
Tamera.McKinniss@act.org

Maria Ofelia Z. San Pedro
ACT, Inc.
sweet.san.pedro@act.org

Aurora Dixon
Michigan State University
rory8891@gmail.com

Jason Way
ACT, Inc.
Jason.Way@act.org

ABSTRACT: There is a growing recognition that non-cognitive skills are relevant to performance in educational and work settings. Given technological advances, researchers have increasingly used educational technologies to explore and affect non-cognitive skills. Educational data mining (EDM) and learning analytics (LA) are key to providing new options for measuring non-cognitive skills in online learning environments. A preliminary study of collaborative problem solving will be used to help illustrate the development of this type of assessment.

Keywords: non-cognitive skills, online learning, assessment, learning analytics

1 KEY NON-COGNITIVE SKILLS

Non-cognitive skills have been referred to by several terms (e.g., behavior, soft skills, social emotional learning, and personality). There is a growing recognition that non-cognitive skills are relevant to performance in educational and work settings (Barrick & Mount, 1991; McAbee, Oswald, & Connelly, 2014; Poropat, 2009; Viswesvaran, Ones & Schmidt, 1996). As defined in Table 1, the key areas found to have the greatest relationship with desired outcomes in education and work include Conscientiousness (Sustaining Effort), Agreeableness (Getting Along with Others), and Emotional Stability (Maintaining Composure).

Table 1. Key Behavioral Domains

Domain	Definition	Common Components
Conscientiousness	The extent to which an individual is careful, disciplined, and achievement oriented	Dependability, order, cautiousness, persistence
Agreeableness	The degree to which a person tends to be kind, considerate, and cooperative, and to focus on interpersonal relationships and social harmony	Cooperation, trust, compassion, altruism

Domain	Definition	Common Components
Emotional Stability	The tendency to respond to stress calmly and manage emotions effectively	Lack of anxiety, personal insecurity, vulnerability to stress

Taken from Camara, O'Connor, Mattern, & Hanson, 2015 see Barrick & Mount, 1991 or Costa & McCrae, 1992 for similar definitions

More specifically, Conscientiousness has consistently been shown to predict achievement from elementary school through college (Poropat, 2009) and it is considered to be the most important behavioral domain related to work (Sackett & Walmsey, 2014). Conscientiousness is considered the next strongest predictor of job performance after cognitive ability (e.g., Almlund, Duckworth, Heckman, & Kautz, 2011; Schmidt & Hunter, 1998). Agreeableness is related to school success both at the K-12 level (e.g., better interpersonal behaviors; Lounsbury, Steel, Loveland, & Gibson, 2004) and in college (e.g., better studying and communication skills; Peterson, Casillas, & Robbins, 2006). In the workplace, agreeableness is related to positive work outcomes, from higher performance (Mount, Barrick, & Stewart, 1998) and interpersonal helping (Gonzalez-Mule, Mount, & Oh, 2014) to greater job satisfaction (Judge & Bono, 2001). Emotional Stability is also related to school success, predicting academic achievement in K-12 (Lounsbury, Gibson, Sundstrom, Wilburn, & Loveland, 2004; Poropat, 2009) and college students (Robbins, Lauver, Le, Davis, & Langley, 2004). In the workplace, emotionality is related to many positive work outcomes, including higher performance (Gonzalez-Mule et al., 2014).

Non-cognitive skills or behaviors can be changed over time. Various programs have been developed to help improve student's non-cognitive skills and subsequently impact their academic outcomes. A recent meta-analysis conducted by Roberts, Luo, Briley, Chow, Su, and Hill (2017) found that interventions can be effective in impacting personality in a clinical setting. This was found to be especially relevant for Emotional Stability. Given the importance of non-cognitive skills, they should be measured and monitored during online learning.

2 NON-COGNITIVE SKILLS AND ONLINE LEARNING

Researchers have used educational technologies to study non-cognitive factors such as academic affect and engagement, both in laboratory settings and in actual classrooms. These advances have progressed in large measure due to the expansion in recent years of the use of computer-based learning environments in schools, providing a rich source of fine-grained data that helps us understand students' learning processes (Ex. ASSISTments, ALEKS, Cognitive Tutor). Fine-grained assessments of cognitive and non-cognitive factors in K-12 education have been shown to predict learning gains (Baker et al., 2010), performance on standardized exams (Pardos et al., 2014), and preparation for future learning (Hershkovitz, Baker, Gowda, & Corbett, 2013).

Through educational data mining (EDM) and learning analytics (LA) (Desmarais & Baker, 2012; Baker & Siemens, 2014), researchers have developed automated models that can infer students' cognitive and non-cognitive factors in real time. Using log data (information recorded by a system) of student interaction with these systems in combination with other data sources (e.g., human observations, sensor data), LA/EDM researchers have developed automated and predictive models of various educational constructs (Baker & Siemens, 2014). These models can infer construct levels in real-time and have found evidence that the inferred constructs are associated with student outcomes. Assessments or measures derived from these models are different from the questionnaire responses and coarse-grained measures or variables (such as demographic information) typically used in research on educational outcomes. Such assessments derived from LA/EDM have been shown to predict educational outcomes such as learning gains (Baker et al., 2004; Cocea et al., 2009; Sabourin et al., 2011) and standardized exam scores (Pardos et al, 2014), and have been widely used in recent years in studying educational phenomena within the context of online learning environments such as intelligent tutoring systems (Baker, D'Mello, Rodrigo, & Graesser, 2010; Walonoski & Heffernan, 2006) and educational games (Shute et al., 2015; Bosch et al., 2015) that produce rich student interaction data.

Information from a preliminary study investigating collaborative problem solving will be used to help illustrate the process of developing and using non-cognitive assessment as part of online learning. The study involved teams of three middle school students working together to solve problems in Physics Playground (Shute & Ventura, 2013), a fun science-based computer game. In order for LA/EDM to be effective, the constructs need to be clearly defined and measured. Therefore, an Evidence Centered Design (ECD) approach was used to define the constructs (e.g., persistence, collaboration, and problem solving), develop the activities and instructions, and ultimately facilitate the measurement of the desired construct indicators through the use of in-game log files as well as out-of-game observed performances (Mislevy, 2011; Mislevy & Haertel, 2006; Mislevy, Steinberg, & Almond, 2002).

3 SUMMARY

It is clear that non-cognitive skills are relevant to online learning. There are several traditional ways to measure these constructs and online learning provides the opportunity to automatically collect additional indicators for these constructs. LA/EDM researchers are and will continue to play a key role in analyzing the wealth of data that can be collected in an online learning environment. More research should continue to be done identifying the indicators most relevant to success in an online learning environment.

REFERENCES

Almlund, M., Duckworth, A., Heckman, J. J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin & L. Woßmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp.1–181). Amsterdam: Elsevier.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004, April). Off-task behavior in the cognitive tutor classroom: when students game the system. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 383-390). ACM.

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223-241.

Baker, R., Siemens, G. (2014) Educational data mining and learning analytics. In Sawyer, K. (Ed.) Cambridge Handbook of the Learning Sciences: 2nd Edition, pp. 253-274.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., et al. (2015). Automatic detection of learning-centered affective states in the wild. In Proceedings of the 2015 International Conference on Intelligent User Interfaces. New York, NY: ACM.

Camara, W., O'Connor, R., Mattern, K., & Hanson, M.A., eds. (2015). *Beyond Academics: A Holistic Framework for Enhancing Education and Workplace Success*. (ACT Research Report 2015-5). Retrieved from http://www.act.org/research/researchers/reports/pdf/ACT_RR2015-4.pdf

Cocea, M., Hershkovitz, A., & Baker, R.S.J.d. (2009). The Impact of Off-task and Gaming Behaviors on Learning: Immediate or Aggregate? *Proceedings of the 14th International Conference on Artificial Intelligence in Education*, 507-514.

Costa, P.T., Jr. & McCrae, R.R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653-665. DOI: 10.1016/0191-8869(92)90236-I.

Desmarais, M.C., Baker, R.S.J.d. (2012) A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments. *User Modeling and User-Adapted Interaction*, 22 (1-2), 9-38.

Gonzalez-Mulé, E., Mount, M. K., & Oh, I. -S. (2014). A meta-analysis of the relationship between general mental ability and nontask performance. *Journal of Applied Psychology*. Journal of Applied Psychology*, 99*(6), 1222-1243*.*

Hershkovitz, A., Baker, R.S.J.d., Gowda, S.M., Corbett, A.T. (2013) Predicting Future Learning Better Using Quantitative Analysis of Moment-by-Moment Learning. Proceedings of the 6th International Conference on Educational Data Mining, 74-81

Lounsbury, J. W., Gibson. L. W., Sundstrom, E., Wilburn, E., & Loveland, J. M. (2004). An empirical investigation of the proposition that' school is work': A comparison of personality-performance correlations in school and work settings. *Journal of Education and Work, 17*, 119–131.

Lounsbury, J. W., Steel, R. P., Loveland, J. M., & Gibson, L. W. (2004). An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence, 33*, 457–466.

McAbee, S. T., Oswald, F. L., & Connelly, B. S. (2014). Bifactor models of personality and college student performance: A broad vs. narrow view. *European Journal of Personality*, *28*, 604-619. doi:10.1002/per.1975

Mislevy, R.J. (2011). Evidence-centered Design for Simulation-based Assessment. Military Medicine. CRESST Report 800. University of California, National Center for Research on Evaluation, Standards, and Student *Testing (CRESST)*, Los Angeles, CA.

Mislevy, R. J., & Haertel, G.D., (2006). Implications of Evidence-Centered Design for Educational Testing. *Educational Measurement: Issues & Practice*, 25, 6-20.

Mislevy, R.J., Steinberg, L.S., & Almond, R. G., (2002). On the roles of task model variables in assessment design. In S. H. Irvine & P. C. Kyllonen (Eds.), *Item Generation for Test Development* (pp. 97-128). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance*, *11*, 145–165.

Pardos, Z. A., Baker, R. S., San Pedro, M., Gowda, S. M., & Gowda, S. M. (2014). Affective States and State Tests: Investigating How Affect and Engagement during the School Year Predict End-of-Year Learning Outcomes. *Journal of Learning Analytics*, 1(1), 107-128.

Peterson, C. H., Casillas, A., & Robbins, S. B. (2006). The Student Readiness Inventory and the Big Five: Examining social desirability and college academic performance. *Personality and Individual Differences*, *41*, 663–673.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135,* 322-338.

Roberts, B. W., Luo, J., Briley, D. A., Chow, P.I., Su, R., & Hill, P.L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, 143(2), 117-141.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261–288. http://doi.org/10.1037/0033-2909.130.2.261

Sabourin, J., Rowe, J., Mott, B., Lester, J. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In Proceedings of Artificial Intelligence in Education 2011, 534-536.

Sackett, P. R., & Walmsley, P. T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science*, *9*, 538–551

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262

Shute, V.J., D'Mello, S., Baker, R., Cho, K., Bosch, N., Ocumpaugh, J., Ventura, M., Almeda, V. (2015) Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. Computers & Education, 86, 224-235.

Shute, V. J. & Ventura, M. (2013). Stealth Assessment: *Measuring and supporting learning in video games*. Cambridge, MA: The MIT Press.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81,* 557–574.

Walonoski, J. A., & Heffernan, N. T. (2006, January). Detection and analysis of off-task gaming behavior in intelligent tutoring systems. In Intelligent Tutoring Systems (pp. 382-391). Springer Berlin Heidelberg.

Exploring Non-Cognitive Reasons behind Success after Failure

Elle Yuan Wang, James Cunningham, Philip Arcuria, Tom Fikes, and Lou Pugliese

Action Lab, Edplus at Arizona State University
{elle.wang; jim.cunningham; parcuria; thomas.fikes; lou.pugliese}@asu.edu

ABSTRACT: The availability of large-scale online courses enabled assessment of academic performances at scale. Equally important is the opportunity to investigate the state of non-cognitive or soft skills of online learners to inform design of interventions to improve learning outcomes. The present analysis looked learners who successfully passed the course, either passed at their first attempts or passed after failed attempts, and compared how their soft skills differ from the rest of learners. The findings suggested that 1) Learners who successfully passed the course tend to have strong skill sets related to time and project management; 2) Learners who successfully passed the course after failed attempts tend to have higher skill sets related to decision making and leadership. Discussions and interpretations are also included.

Keywords: Online learning, non-cognitive skills, learning analytics

1 INTRODUCTION

Pathways to success vary. In the context of academic settings, some students may experience more failures than others, before achieving success. Some failures may be productive and necessary (Kapur, 2008; Kapur, 2010). Learners with varied prior knowledge may require a diverse range of "desirable difficulties" (Schmidt & Bjork, 1992). Nevertheless, learners may treat failure differently. Some learners may be more discouraged than others when facing failure, which may be due to cognitive and non-cognitive differences among learners. While educators and researchers are developing various cognitive interventions to help learners succeed, more studies have looked into decoding the non-cognitive differences among learners (McAbee, Oswald, & Connelly, 2014; Farkas, 2003).

Researchers have demonstrated that, on average, every dollar invested in development of non-cognitive skillsets such as those on social emotional learning programs yields over 10 times that amount in long-term benefits (Belfield et al., 2015).

Social and emotional learning constitutes a critical part of non-cognitive skillsets. Non-cognitive skills (Heckman & Rubinstein, 2001) refer to social and emotional learning capabilities beyond academic knowledge. These skills comprise a diverse range of aspects including conscientiousness (Camara et al., 2015), growth mindset (Dweck, 2006), time management, and self-regulation (Gutman & Schoon, 2013), to name a few.

It has been shown that the predictive power of non-cognitive skills on academic achievement across wide range of educational settings is at least equal to or better than the predictive power of cognitive skills. Non-cognitive skills have been associated with various academic outcomes (e.g., Lleras, 2008) including persistence in postsecondary settings with attendance and retention (e.g., Credé, Roch, & Kieszezynka, 2010) and engagement (McClenney, Marti, & Adkins, 2006). Moreover, non-cognitive skills have also been linked to career advancement (Heckman et al., 2006), well-being (Strickhouser, Zell, & Krizan, 2017) and key 21st century competencies, such as critical thinking and problem solving (Buckingham Shum & Crick, 2016).

Therefore, it is imperative to understand why some learners are more likely to continue their learning processes after failure than others. To respond to this research question, the present study included analyses to investigate how learners who succeeded in passing an online mathematics course after failed attempts differ regarding to their non-cognitive skillsets.

2 METHOD

2.1 Data Source/Participants

This study used the passing rates of online college algebra students at a large public university in the southwestern United States to examine the relationship between soft skills and academic performance. The online course was offered during the fall and spring of the 2016-2017 academic school year and used the adaptive learning system ALEKS (e.g., Hagerty & Smith, 2005) which guides students through practice and assessment of 383 mathematics topics related to college algebra. A passing grade in this course is considered to be a letter grade of “C” or better. Students in this course had an opportunity to retake college algebra at no additional cost if they were unsuccessful in passing on the first try. Although, an instructor is assigned to each section of the online course, the students’ primary interaction is with the adaptive online learning system as they read about each math concept, work through problems, request examples, and take assessments.

2.2 25 soft skill competency assessments

A total of 255 learners of the math course voluntarily participated in a non-cognitive skillsets assessment developed by Indigo Project (Indigo, 2017), designed as part of an online orientation course. 25 types of non-academic skillsets (Table 1), developed based on the 21st century competencies (TTI, 2012), were collected on the student level. The 25 types of skillsets were developed based on the 21st century competencies. This type of assessment has been used in the field of engineering entrepreneurship education (Pistrui et al., 2011), surgical training (Bell et al., 2012), as well as secondary science education (Bonnstetter, 2003).

2.3 2 Types of outcomes

Two types of binary outcome measures were included in the present analyses: “Passing the course” and “Passing the course after failure”. “Passing the course” was computed as 1 for passing; the rest were coded as 0. Similarly, “Passing the course after failure” was computed as 1 for learners who passed the course after experiencing at least one failed attempt; the rest were coded as 0.

2.4 Analyses

A principal component analysis based on the 25 competency assessments was conducted. The extracted principal components were then used in two logistic regression models to predict the two types of outcome measures: passing the course, and passing the course after failed attempts.

3 RESULTS

3.1 Principal Component Analyses

A principle component analysis of the 25 skillsets was conducted. Five components explaining 63.76% of the variance were extracted. An oblimin rotation provided the best-defined component structure. Loadings higher than .45 are in bold in Table 1 to highlight the stronger contributions from the skillsets for each of the 5 components.

Component 1: High on Empathy and Teamwork Skills
The first principal component is strongly correlated with five of the original variables. The first principal component grows with increasing “Appreciating Others”, “Customer Focus”, “Diplomacy”, Interpersonal Skills:”, “Teamwork”, and “Understanding Others” scores.

Component 2: High on Decision Making and Organizational Skills
The second principal component is strongly correlated with three of the original variables. This component grows with increasing “Decision Making”, “Planning Organizing”, and “Project Management”.

Component 3: High on Creativity
The third principal component is strongly correlated with five of the original variables. This component grows with increasing “Conceptual Thinking”, “Continuous Learning”, “Creativity Innovation”, “Flexibility”, and “Futuristic Thinking” scores.

Component 4: Low on Leadership Skills
The fourth principal component is strongly correlated with four of the original variables. This component grows with decreasing “Conflict Management”, “Influencing Others”, “Leadership”, and “Negotiation” scores.

Component 5: Low on Time management Skills

The fifth principal component is strongly correlated with five of the original variables. This component grows with decreasing "Goal Orientation", "Project Management", "Resiliency", "Self-Starting", and "Time Management" scores.

Table 1: Factor loadings based on a principle component analysis with oblimin rotation for 25 skillsets from the Indigo assessment.

	Components				
	1	2	3	4	5
Appreciating Others	**.854**	-.084	.134	.143	.111
Conceptual Thinking	.074	.038	**.864**	.041	.142
Conflict Management	.421	.205	-.016	**-.575**	.054
Continuous Learning	-.061	.232	**.588**	-.079	-.215
Creativity Innovation	.011	-.160	**.803**	-.086	-.049
Customer Focus	**.556**	.268	.011	.046	-.107
Decision Making	-.007	**.874**	.050	-.022	.090
Diplomacy	**.597**	.258	-.001	-.137	.058
Development Coaching	.444	.123	-.092	-.260	-.303
Flexibility	.130	-.007	**.595**	.102	-.360
Futuristic Thinking	-.052	.112	**.745**	-.163	.064
Goal Orientation	.006	.047	.135	-.244	**-.653**
Influencing Others	-.074	-.088	.179	**-.835**	-.096
Interpersonal Skills	**.594**	-.015	.104	-.327	-.077
Leadership	.054	-.111	.186	**-.480**	-.434
Negotiation	.214	.241	.087	**-.623**	.045
Personal Accountability	.345	.151	.240	.290	-.415
Planning Organizing	.107	**.682**	-.088	.008	-.228
Problem Solving	-.056	**.717**	.126	-.031	-.059
Project Management	.008	.331	.042	-.254	-.473
Resiliency	.208	.037	.148	.057	**-.502**
Self-Starting	.165	-.242	.163	-.082	**-.585**
Teamwork	**.674**	-.155	-.153	-.121	-.164
Time Management	-.066	.251	-.157	.006	**-.788**
Understanding Others	**.621**	-.085	.228	-.142	-.109

3.2 Logistic Regression

Logistic regression analyses were conducted to test whether the five components derived from the 25 skillsets can predict learners' performance. Two types of indicators for performances were used: passing the course and passing the course after failure.

3.2.1 Passing the course

A logistic regression using the backward selection method was conducted to predict whether a learner passes the course based on the 5 component scores derived from the 25 skillsets. The final model included Component 1 and Component 5. The Wald test results showed that Component 1, *High* on Empath and Teamwork Skills, χ^2 (1) = 4.951, p = .026, and Component 5, *Low* on Time Management Skills, χ^2 (1) = 5.793, p = .016 were individually each statistically significant within the combined model.

Increasing Component 1 was associated with a reduction in the likelihood of passing the course; increasing Component 5 was also associated with a reduction in the likelihood of passing the course. In other words, increasing scores on empath and team work skills were associated with a lower likelihood of passing the course; whereas increasing scores on time management scores were associated with a higher likelihood of passing the course.

Table 2: Logistic Regression Analysis on Passing the Course

	B	S.E.	Wald	df	Sig.	Exp(B)
Component 1	-.344	.155	4.951	1	**.026**	.709
Component 5	-.372	.154	5.793	1	**.016**	.690
Constant	.083	.136	.369	1	.543	1.086

3.2.2 Passing after failure

A logistic regression using the backward selection method was conducted to predict whether a learner passes the course after prior failures based on the 5 component scores derived from the 25 skillsets. The final model included Component 2, Component 3, and Component 4. The Wald test results showed that Component 2, *High* on Decision Making and Organizational Skills, χ^2 (1) = 5.109, p = .024, and Component 4, *Low* on Leadership Skills, χ^2 (1) = 9.800, p = .002 were individually each statistically significant within the combined model.

Increasing Component 2 was associated with an increase of likelihood of passing the course after failure; increasing Component 4 was associated with a reduction in the likelihood of passing the course after failure. In other words, increasing scores on decision making and organizational skills as well as leadership skills were associated with higher likelihood of passing the course after failure.

Table 3: Logistic Regression Analysis on Passing after Failure

	B	S.E.	Wald	df	Sig.	Exp(B)
Component 2	.847	.375	5.109	1	**.024**	2.332
Component 3	-.634	.346	3.353	1	.067	.531
Component 4	-1.410	.450	9.800	1	**.002**	.244
Constant	-3.822	.540	50.165	1	.000	.022

4 DISCUSSION

4.1 What predicts passing?

Based on the results above, higher scores on empathy and teamwork were negatively associated with the probability of passing the course. This result is somewhat counter-intuitive. One possible interpretation is that data for the present analysis was collected from a math course where the final grade was not dependent upon interaction with fellow learners. For this specific course, it is possible that learners may benefit more from focused independent learning sessions.

By contrast, the time management skillsets are positively associated with passing the course. This finding is expected since time management skills were found to be an important factor that influenced the learners' online learning experiences (e.g., Song, et al., 2004). This finding suggests that improvement on online learners' time management skills may lead to better learning outcomes.

4.2 What predicts passing after failure?

Planning, organizational, and leadership skills were found to positively associate with the likelihood of passing the course after failure. It is possible that learners who are more skilled in planning and organization are more likely to objectively treat their failure as a useful learning process and decided to keep trying after that. It is worth investigating as to why this group of learners failed the course before they achieved success afterwards.

5. CONCLUSION

The present paper investigated non-cognitive differences behind learners who took different pathways toward success in an online mathematics course. Results from this paper suggested that it is meaningful to investigate differences in terms of learners' soft skill competencies to help understand why learners exhibit differed learning outcomes. Future studies, by employing follow-up interviews with learners, can look into specific reasons behind those identified soft skills. For example, one follow-up research question could be: Do learners whose assessment showed a lack of time management skillsets actually share the same concern when interviewed? If so, what are the reasons behind the lack of the time management skills? This way, targeted and effective interventions can be designed to help learners to achieve better learning outcomes.

REFERENCES

Bell, R. M., Fann, S. A., Morrison, J. E., & Lisk, J. R. (2012). Determining personal talents and behavioral styles of applicants to surgical training: a new look at an old problem, part II. *Journal of surgical education*, *69*(1), 23-29.

Belfield, C., Bowden, A. B., Klapp, A., Levin, H., Shand, R., & Zander, S. (2015). The economic value of social and emotional learning. *Journal of Benefit-Cost Analysis*, *6*(3), 508-544.

Bonnstetter, R. J. (2003). A triad of disposition instruments used in secondary science education to help teachers better understand self and others. *Science Education International*.

Buckingham Shum, S., & Crick, R. D. (2016). Learning Analytics for 21st Century Competencies. *Journal of Learning Analytics*, *3*(2), 6-21.

Camara, W., O'Connor, R, Mattern, K, Hanson, M. A. (2015). Beyond Academics: A Holistic Framework for Enhancing Education and Workplace Success. (ACT Research Report 2015-4). Retrieved from http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-4.pdf

Credé, M., Roch, S. G., & Kieszczynka, U. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *80*, 272–295.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House Incorporated.

Hagerty, G., & Smith, S. (2005). Using the web-based interactive software ALEKS to enhance college algebra. *Mathematics and Computer Education*, *39*(3), 183.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*, *91*(2), 145-149.

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. Social Science Research, 37(3), 888-902.

McClenney, K., Marti, C. N., & Adkins, C. (2006). *Student engagement and student outcomes: Key findings from CCSSE validation research.* Austin, TX: Center for Community College Student Engagement.

Indigo. (2017). Assessment validity, reliability, measurement variables, competitor landscape, customers & partners. Retrieved from: https://static1.squarespace.com/static/56c15de859827ef22c082045/t/5977c46220099e603532e532/1501021287959/Indigo+Validity+Datapoints+Competitors+Schools.pdf

Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. *Education Empowerment Foundations, London*. Retrieved from: http://hdl.voced.edu.au/10707/287500

Kapur, M. (2008). Productive failure. *Cognition and instruction*, *26*(3), 379-424.

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, *38*(6), 523-550.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217.

Silva, E. (2009). Measuring skills for 21st-century learning. *Phi Delta Kappan*, *90*(9), 630-634.

Song, L., Singleton, E. S., Hill, J. R., & Koh, M. H. (2004). Improving online learning: Student perceptions of useful and challenging characteristics. *The internet and higher education*, *7*(1), 59-70.

Strickhouser, J.E., Zell,E., & Krizan, Z. (2017). Does personality predict health and well-being? A metasynthesis. *Health Psychology,* 36(8), 797-810.

Pistrui, D., Bonnstetter, R., Bonnstetter, B. J., & Fry, C. C. (2011). Creating, Educating and Assessing a New Class of Entrepreneurial Minded Engineers. *The Journal of Engineering Entrepreneurship*, *2*(2), 1-14.

TTI Success Insights. (2012). *TTI Technical Reports Compendium*. Retrieved from: https://static1.squarespace.com/static/56c15de859827ef22c082045/t/59789c569f74560e1d74da91/1501076569619/TTI+Technical+Report.pdf

Non-cognitive Assessments at Scale: MOOCs and Employability

Tawanna Dillahunt
University of Michigan
tdillahu@umich.edu

Yuan Wang
EdPlus at Arizona State University
elle.wang@asu.edu

ABSTRACT: Online learning environments such as Massive Open Online Courses (MOOCs) have been shown to support employability among learners, and to provide staff with development opportunities for their employees. However, while certain aspects of human capital can be properly assessed in MOOCs, there are no formal assessments to evaluate all dimensions of employability: career identity, personal adaptability, and social and human capital—which can lead to successful employment outcomes. We propose Fugate et al. psycho-social construct of employability to provoke a discussion about whether and how these dimensions can be used to inform the assessment of non-cognitive skills in online learning environments. Employability is vital to reaching online learning environments' potential to support those learners who are motivated to take online courses such as MOOCs for employment-related reasons.

Keywords: MOOCs, employability, employment, social capital, non-cognitive skills

1 INTRODUCTION

Massive Open Online Courses (MOOCs) provide Internet-connected learners with opportunities to take a wide range of courses from some of the world's most elite schools (Nanfito, 2014); MOOCs also provide opportunities to connect learners to other learners around the world (Kulkarni, Cambre, Kotturi, Bernstein and Klemmer, 2015). In addition, MOOCs help to remove several barriers associated with traditional education such as costs and are regularly updated with new material (Nanfito, 2014). Given these paybacks, MOOCs and other learning environments are in a position to support education and employability. Yet, the vast majority of research exploring MOOCs investigates enrollment, user populations, participation, completion and retention rates, platform development, pedagogy, motivations, business models, and social network analyses (e.g., Clow, 2013; Cooper and Sahami, 2013; Dellarocas and Van Alstyne, 2013; Dillahunt, Chen, and Teasley, 2014; Dillahunt, Wang, and Teasley, 2014; Ho, Reich, Nesterko, Seaton, Mullaney, Waldo and Chuang, 2014; Kasunic, Hammer, and Ogan, 2015; Kizilcec and Schneider, 2015; Kizilcec, Piech and Schneider 2013; Kolowich, 2013; Rosé and Siemens; Rosé, Carolson, Yang, Wen, Resnick, Goldman, Sherer, 2014). However, understanding MOOCs' roles and relationship with post-course employment is still relatively unexplored (Calonge and Shah, 2016; Dillahunt, Ng, Fiesta, and Wang, 2016).

1.1 The Employability Framework

Individuals must be flexible and adaptable to manage rapidly changing career landscapes (Fugate, Kinicki, and Ashforth, 2004) that are becoming increasingly common today. A person's willingness and ability to adapt is critical to having a successful career (Hall, 2002). Fugate et al. propose employability, a "psycho-social construct that embodies individual characteristics that foster adaptive cognition, behavior, and affect, and enhance the individual-work interface" (p. 15, 2004). Employability consists of three dimensions: career identity, personal adaptability, and social and human capital. We then propose these three aspects of this construct be used as a non-cognitive assessment in online learning environments such as MOOCs.

Career identity is an individual's definition of self in the career context. It describes who the person is or who the person wants to be and can operate as a "cognitive compass" motivation an individual to create or realize opportunities. Personal adaptability is one's ability to adapt to changing situation by changing the following personal factors: one's propensity to learn, optimism, openness, generalized self-efficacy and one's internal locus of control (Fugate, Kinicki, and Ashforth, 2004). Social and human capital are embedded into one's career identity. Social capital is described as the benevolence or goodwill intrinsic to one's social networks and it is critical for achieving occupational goals. Human capital consists of several factors such as education and experience, which is the strongest predictor of career progression (Judge, Cable, Boudreau, and Bretz, 1995), emotional intelligence (Wong and Law, 2002), and cognitive ability (Tharenou, 1997). According to Fugate et al. "one's perceived ability to identify and realize career opportunities is derived from their career identity, personal adaptability, and social and human capital" (p. 26, 2004).

1.2 MOOCs and Employability

In a systematic qualitative literature review of sixteen education and technology-enhanced articles that identified MOOCs' potential to bridge the skills gap between employers and students, researchers found that MOOCs potential to help students attain relevant skills before employment is relatively unexplored (Calonge and Shah, 2016). In addition, the authors suggest that MOOCs are global stakeholders in increasing opportunities for corporations to provide staff training and development to their employees as well as to those job seekers looking for employment. A qualitative study of whether MOOCs are platforms for employability revealed that while MOOCs support some aspects of human capital, they provide little support for career identity, personal adaptability, or social capital (Dillahunt, Ng, Fiesta, Wang, 2015).

2 LOOKING AHEAD: ASSESSMENTS OF EMPLOYABILITY

While assessment and support for the vocational and cognitive aspects of human capital is prevalent in these online learning environments, little is known about non-cognitive aspects such as one's ability to acquire social capital in MOOCs, a learner's career identity and personal adaptability over time while taking online courses. We argue that the employability framework can inform assessments of these non-cognitive skills especially in the context of online learning environments such as MOOCs. These platforms provide the perfect space to assess employability dimensions as a large percentage of learners are employees who wish to transition to new careers (Dillahunt, Ng, Fiesta, Wang, 2015; Kizilcec and Schneider, 2015) and employability is especially beneficial to employees in transition who may be coping with job search and even job loss (Fugate, Kinicki, and Ashforth, 2004). In addition, Fugate et al. argue that more research is needed to: 1) further define the construct of employability—what constitutes low versus high employability? and 2) what role does employability play in various work-related phenomena (2004)? Therefore, we propose to investigate the relationship between various non-cognitive skills and employability to help answer these questions.

REFERENCES

Calonge, D. S., & Shah, M. A. (2016). MOOCs, Graduate Skills Gaps, and Employability: A Qualitative Systematic Review of the Literature. *The International Review of Research in Open and Distributed Learning*, *17*(5).

Clow, D. (2013). MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 185-189). Leuven, Belgium: ACM.

Cooper, S., & Sahami, M. (2013). Reflections on Stanford's MOOCs. *Communications of the ACM, 56*(2), 28-30.

Dellarocas, C., & Van Alstyne, M. (2013). Money models for MOOCs. *Communications of the ACM, 56*(8), 25-28.

Dillahunt, T., Chen, B., & Teasley, S. (2014, March). Model thinking: demographics and performance of MOOC students unable to afford a formal education. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 145-146). ACM.

Dillahunt, T. R., Wang, B. Z., & Teasley, S. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in Open and Distributed Learning*, *15*(5).

Dillahunt, T. R., Ng, S., Fiesta, M., & Wang, Z. (2016, February). Do Massive Open Online Course Platforms Support Employability?. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 233-244). ACM.

Fugate, M., Kinicki, A.J., & Ashforth, B.E. (2004). Employability: A psycho-social construct, its dimensions, and applications. Journal of Vocational Behavior, 65, 14–38.

Ho, A. D., Reich, J., Nesterko, S. O., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITx: The First Year of Open Online Courses (HarvardX and MITx Working Paper No. 1). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2381263

Kasunic, A., Hammer, J. & Ogan, A. (2015). Cultural Relevance in MOOCs: Asking About Socioeconomic Context. In Proceedings of the Second (2015) ACM Conference on Learning@ Scale (pp. 389-392). ACM.

Kizilcec, R.F. and Schneider, E. (2015). Motivation as a lens to understand online learners: Towards data-driven design with the olei scale. ACM TOCHI 22(2), 24.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). *Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses.* Paper presented at the Third International Conference on Learning Analytics and Knowledge LAK '13, Leuven, Belgium.

Kolowich, S. (2013, April 29). Why some colleges are saying no to MOOC deals, at least for now. The Chronicle of Higher Education. Retrieved from http://chronicle.com/article/Why-Some-Colleges-Are- Saying/138863/

Kulkarni, C., Cambre, J., Kotturi, Y., Bernstein, M.S., and Klemmer, S.R. 2015. Talkabout: Making Distance Matter with Small Groups in Massive Classes. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '15). ACM, New York, NY, USA, 1116-1128. DOI: http://dx.doi.org/10.1145/2675133.2675166

Nanfito, M. (2014). MOOCs: Opportunities, impacts, and challenges: massive open online courses in colleges and universities. CreateSpace Independent Publishing Platform. Lexington, KY.

Rosé, C. P., & Siemens, G. (2014, October). Shared task on prediction of dropout over time in massively open online courses. In *Proc. of EMNLP* (Vol. 14, p. 39).

Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., & Sherer, J. (2014, March). Social factors that contribute to attrition in MOOCs. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 197-198). ACM.

Schuller, T. (2007) Reflections on the use of social capital, Review of Social Economy, 65:1, 11-28, DOI: 10.1080/00346760601132162

Itero: A New Tool for Exploring Writing Self-efficacy through Revision History Analytics

Selen Turkay
Queensland University of Technology
selen.turkay@qut.edu.au

Daniel Seaton & Andrew Ang
Harvard University
daniel_seaton@harvard.edu & andrew_ang@harvard.edu

ABSTRACT: The goal of this presentation is to contribute to the discussion on the assessment of non-cognitive skills in digital environments using writing revision history analytics. We will share the results from an exploratory pilot study with a writing revision analytics tool, called Itero. Writing is a fundamental part of student development and academic assessment. Studies indicate that writing proficiency predicts academic success, helps development of critical thinking skills, and is of great importance after formal education in professional and personal settings. Writing self-efficacy is a non-cognitive skill, critical for writing performance and persistence. However, scaling the measurement of self-efficacy presents a challenge: self-report surveys are the most commonly used methods. In this short paper, we present findings from a pilot study with college students where we found evidence that Itero can make student writers more aware of their writing process, which may influence their self-regulatory processes. We also present evidence that may help understand whether writing revision analytics can shed light on which students have low or high self-efficacy from their writing behavior. We discuss how Itero may be used to increase students' writing self-efficacy by tapping into mastery and social comparison information resources, proposed by Bandura (1977).

Keywords: writing self-efficacy; writing revision analytics; essay writing

1 INTRODUCTION

Writing is a critical skill for success in schools and beyond. Students need to communicate and express their ideas through written assignments across a majority of academic fields. Writing skills are also an integral part of most jobs. Yet, many students are not able to adequately write extended texts, and writing self-efficacy was proposed to be one of the reasons for this. Self-perception of ability, or self-efficacy, is one of the non-cognitive skills that moderate students' academic performance (Gutman & Schoon, 2013). Self-efficacy affects people's cognitions, motivations, affective processes and ultimately their behavior (Bandura, 1977). Students with the same level of writing skills might perform differently depending on their belief of efficacy in writing (Bandura, 1977). Bruning, Dempsey, Kauffman, McKim, and Zumbrunn (2013) argue that writing-self efficacy is critical because writing tasks are demanding and motivational conditions are less than ideal.

Strategies and tools that enhance students' self-efficacy during writing revision would be a significant contribution to students' academic development (Pajares, 2003). Yet, writing self-efficacy is not easy to measure. The most common method of measuring self-efficacy is survey. Considering the low completion rates of surveys by students, we propose and explore using writing analytics and visualizations to measure and, ultimately, increase students' self-efficacy.

Recent studies have shown evidence that analytics have affordances to help give writers meaning-making opportunities to take action (Vatrapu, Reimann, Bull, & Johnson, 2013; Durall, & Gros, 2014; Tabuenca, Kalz, Drachsler, & Specht, 2015). The work exploring the relationship between writing analytics and self-regulatory processes to improve self-efficacy is still in its infancy. Studies, to date, explored designing dashboards to support reflection and self-monitoring (McNely, Gestwicki, Hill, Parli-Horne, & Johnson, 2012; Vatrapu et al., 2013) without in-depth exploration of the relationship between user characteristics and patterns in analytics.

According to Bandura (1977), students form self-efficacy beliefs by interpreting information from their environment. The most powerful source of information is previous performance (Pajares, Johnson & Usher, 2007). Other influential sources of information come from vicarious experience from observation and social comparison, and from social persuasion. Students can also develop self-beliefs through experiencing physiological and emotional states, such as exhilaration and anxiety, (Bandura, 1977; Pajares et al., 2007). We have developed an application that aims to present mastery and social comparison information that can potentially reinforce students' self-efficacy.

2 ITERO

Itero is an application that shows detailed revision history for writers to observe their writing behavior. This is achieved by using writing analytics and visualizations, building upon prior projects and studies in the field (e.g., Wang, Olson, Zhang, Nguyen, & Olson, 2015). It can be used to understand the temporal nature of revision patterns and how those relate to activities within specific passages of a written text. Such patterns can be used to identify writing strategies that would potentially lead to computer-aided interventions, e.g., automatically asking a student to pause and reflect on their writing of a paragraph.

Itero uses detailed revision history from Google Docs to enable users to further analyze and visualize the writing process. After authenticating with their Google user account, and giving permission for Itero to access their Google Drive files, users can browse a list of Docs in their Drive that they own. They may select a particular document(s) to view in Itero - at which point Itero retrieves the detailed revision data for that particular document from Google Drive. In its current design, Itero has three main affordances:

- Individual's writing analytics: visualizing temporal and spatial patterns of the writing process (potential to improve mastery);

- Collaborative writing analytics: visualizing the contributions by different co-writers (facilitate for social-comparison);
- Replay: character-by-character replay of the writing process in different play speeds (close self-monitoring).

The goal of this exploratory study was two-fold: 1) investigate the potential role of Itero on individuals' understanding of their own writing revision process during a writing task (e.g., can participants interpret information presented by Itero?); 2) examine the relationship between writing analytics patterns and self-efficacy beliefs. Because this is the first study we conducted with Itero, we employed a mixed method approach, combining surveys with semi-structured interviews and writing analytics to explore the experiences of potential users. Before we further developed our application, we wanted to understand the barriers faced by users when using Itero during an individual writing task, and learn about their needs.

2.1 Participants

We recruited nine undergraduate students (5 female and 4 male) in a large private university in the east coast of the United States. Students received $15 for one hour of their time. The study took place in a usability research laboratory of the university on a 1-1 basis. All participants were familiar with Google docs but did not use the application intensively in their academic work. On average, participants rated their writing fluency as very fluent.

2.2 Procedure and Instruments

After signing the informed consent form, we conducted a 5-minute interview to learn about participants' use of different writing tools and their perceived needs to be better writers. Once seated in front of the lab PC, participants filled out a short survey. They were then given the following writing prompt and were asked to write a two pages long essay supporting one of the beliefs:

> Some people believe that it is imperative for individuals living in developed nations to reduce their energy consumptions and lead a more sustainable lifestyle, given the evidence for global climate change. Others believe that such drastic lifestyle changes are unwarranted, based on the existing evidence for global climate change.

After 25 minutes of writing, participants were asked to look at the visuals in Itero and study them for couple of minutes. After, they continued with their writing for 10 more minutes. After completing the writing, we conducted a short interview with participants. We used the multifactor scale developed by Brunning et al. (2013) to measure participants' self-efficacy for writing. The factors are ideation, conventions, and self-regulation.

3 FINDINGS

3.1 User Experiences with Itero

Interview data revealed that, overall, participants found Itero as a novel application. As one of the participants put it they "had never seen anything like this before." However, novelty does not justify adoption of a tool. When asked whether they would use Itero for their academic writing, all nine of the participants expressed doubt in using it for individual work, but they said they would use it for collaborative writing and group projects. They proposed two main reasons for it: 1) to know how much each co-writer contributes (e.g., quantity, accountability, social comparison), and 2) to understand who contributes which ideas (e.g., quality/content). In other words, participants wanted to see how engaged their co-writers are. As one of the participants, P3, put it she would use Itero "to see where ideas came from and to see how engaged every single member of the group were in the writing." P4 indicated that he would use Itero if an assignment has participation grade because Itero would allow him to gauge how much he is contributing compared to his teammates.

For individual writing process, P7 stated that Itero can help her understand her writing and revision process. She said "…tracking how I write, to notice what I've worked the most on, moved around the most, etc. could help figure out what parts of an essay I've thought through or drafted more and which parts I've left as they stand." She further stated that "In the context of a normal assignment, you might not get much of a chance to use it. It would be useful to help develop writing skills."

All participants found the replay function as potentially the most useful for their writing. P2 said replay would help her to find out "…if there is a section you keep going back while writing. Maybe the reason is you don't know if the paragraph fits because you keep questioning it." P5 put it as "Seeing your document perform…" Monitoring their document perform may influence students' emotional states, which we will investigate in future studies.

3.2 Writing Patterns

We detected different writing patterns among participants. Participants who scored high on self-efficacy wrote considerably longer essays. These participants with higher self-efficacy also showed different patterns than those who had lower self-efficacy. For instance, the top two images in Figure 1 belong to participants with the lowest self-efficacy and the bottom two below to those who with

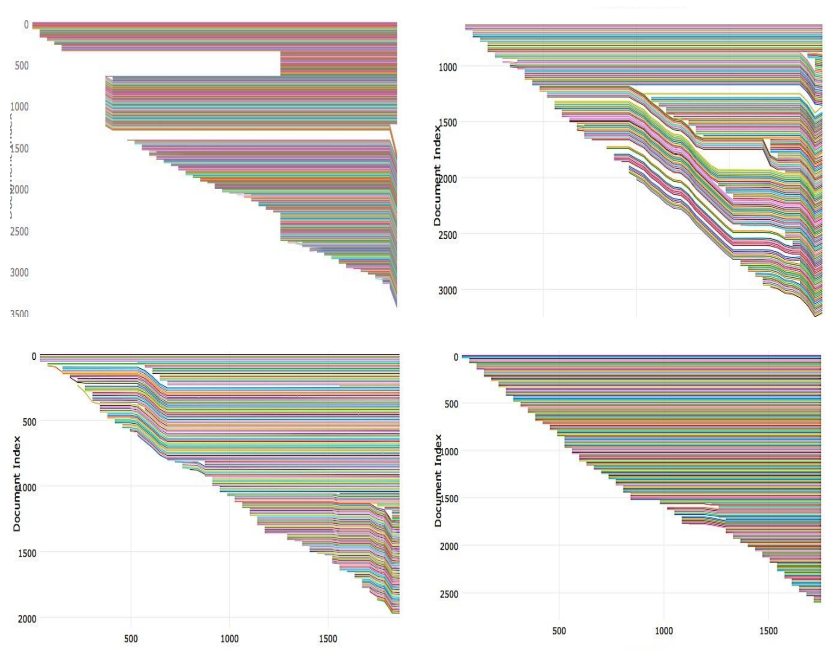


Figure 1: Visualizations showing data tracking with lowest (top row) and highest self-efficacy (bottom row). Top left shows many copy-paste behavior (represented by large blocks)

the highest self-efficacy. Participants with the lowest self-efficacy also scored the two lowest grade level categories (P9: 10 to 11th grade and P5: 8th to 9th grade) based on Flesh-Kinkaid text standard. Both participants (P6, P7) who had high self-efficacy scored 13th to 14th grade level text standard (3rd highest among all participants).

4 DISCUSSION

In this short paper, we presented an exploratory study on how Itero's features may help us discover students' self-efficacy beliefs. Interpreting the information from different sources is an important process for students' self-efficacy development. Since Itero presents a wide variety of information, one of our goals was to explore whether these were comprehensible to students. We invited participants to observe their writing process using multiple features of Itero, and report their experiences. Overall, participants responded to Itero positively. They became aware of writing behaviors such as focusing too much on a single section or not paying attention to grammar. We also learned that participants found some information difficult to interpret.

We also found evidence of a relationship between participants' self-efficacy ratings and their writing revision behaviors. Students with the lowest self-efficacy used copy-pasting more often and deleted large sections of their writing. We only focused on participants with the highest and the lowest self-efficacy. In our future study, we will investigate what kind of patterns may emerge from students with somewhat high or low writing self-efficacy. The ultimate goal is to train algorithms to detect students' writing self-efficacy automatically so that we can utilize computer-aided interventions to improve students' writing self-efficacies, e.g., using motivational messages, or social persuasion.

Due to the short nature of the study, we were unable to study how students may observe their mastery development using visuals of Itero. Our next study will use a control group, and will take place in an online setting with a hundred participants over two weeks to increase the ecological fidelity of the experimentation. Experimental setup will enable us to investigate the impact of Itero on writers' self-efficacy. Using data from these participants, we will also be able to further study the relationship between self-efficacy and patterns in writing revision behaviors.

We look forward to discussing with workshop participants and organizers the potential uses of Itero to measure and improve students' writing self-efficacies in online distributed learning environments.

REFERENCES

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, *105*(1), 25–38.

Durall, E., & Gros, B. (2014). Learning Analytics as a Metacognitive Tool. In *CSEDU (1)* (pp. 380–384).

Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. *Education Empowerment Foundations, London*.

McNely, B. J., Gestwicki, P., Hill, J. H., Parli-Horne, P., & Johnson, E. (2012). Learning Analytics for Collaborative Writing: A Prototype and Case Study. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (pp. 222–225). New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330654

Pajares, F. (2003). Self-Efficacy Beliefs, Motivation, and Achievement in Writing: A Review of the Literature. *Reading & Writing Quarterly*, *19*(2), 139–158.

Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. Research in the Teaching of English, 104-120.

Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers & Education*, *89*, 53–74.

Vatrapu, R., Reimann, P., Bull, S., & Johnson, M. (2013). An Eye-tracking Study of Notational, Informational, and Emotional Aspects of Learning Analytics Representations. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 125–134). New York, NY, USA: ACM.

Wang, D., Olson, J. S., Zhang, J., Nguyen, T., & Olson, G. M. (2015). DocuViz: Visualizing Collaborative Writing (pp. 1865–1874). ACM Press.

Enhancing global engagement through environmental education MOOCs

Yue Li
Cornell University
yueli@cornell.edu

Marianne Krasny
Cornell University
mek2@cornell.edu

ABSTRACT: Massive Open Online Course (MOOCs) can provide a platform for participants to connect and share information with each other, yet little information is available on educators' global engaged learning experience. Based on a series of MOOCs for professional development in environmental education, this research applies the Global Engagement Survey to measure MOOC participants' global experience including intercultural competence and civic engagement. The results provide insights into measuring MOOC participants' global learning, as well as suggest ways of designing and teaching MOOCs to foster participants' understanding of international perspectives and actions as global citizens in environmental education.

Keywords: MOOCs, global engagement, intercultural competence, civic engagement, environmental education

1 INTRODUCTION

Massive Open Online Courses (MOOCs) can be considered as a form of social learning (Bandura, 1977; Wals, 2007) by providing an opportunity for participants to exchange ideas and learn from each other (Brinton et al., 2014). Through such a social learning environment, MOOCs address not only knowledge access and aquisition, but also help participants develop a variety of non-cognitive skills (Hollands & Tirthali, 2014). By engaging participants from multiple countries, MOOCs have the potential to increase participants' intercultural competence such as communication skills and self-awareness. Our environmental education MOOCs in particular emphasize civic engagement, which could be enhanced by a global learning experience.

MOOC research on participants' online interactions has focused on participants' engagement within the course (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014) and communication networks (Kellogg, Booth, & Oliver, 2014). MOOC research also has examined the relationship between participants' cultural background with their learning performance (Liu et al.) and explored ways to enhance achievement for participants from different countries (Kizilcec, Saltarelli, Reich, & Cohen, 2017). Little is known about how MOOCs increase communication skills, intercultural awareness, and civic engagement through interactions among participants with diverse cultural background.

The Global Engagement Survey (GES) (Reynolds & Hartman, 2016) was developed based on established surveys such as the International Volunteering Impacts Survey (Lough, McBride, & Sherraden, 2009), Global Citizenship Scale (Morais & Ogden, 2011), and Global Perspective Inventory (Braskamp, Braskamp, Merrill, & Engberg, 2008). It seeks to examine university students' global engagement through global service (Hartman & Kiely, 2014) such as study abroad (Paige, Fry, Stallman, Josić, & Jon, 2009). By adapting the GES to MOOCs, we are able to measure MOOC participants' global engagement as a non-cognitive outcome of an online learning experience.

To enhance our understanding of participants' global engaged learning through environmental education MOOCs, we adapted the GES for use in multiple environmental education courses. The specific research questions are: 1) How do MOOC participants enhance their intercultural competence? 2) How do MOOC participants increase their civic engagement?

2 METHODS

2.1 Participants

The MOOCs for professional development in environmental education that are the focus of this study include: *Introduction to Environmental Education*, *Environmental Education Outcomes*, *Urban Environmental Education*, and *Global Environmental Education*. We offered the courses through EdX Edge and Canvas used social media (Facebook and WeChat) for exchanging ideas and resources and facilitating discussions among MOOC participants. Each course asked participants individually or collaboratively to complete a project to design a lesson plan or develop a case study to address their local environmental challenges. About 550 participants from 50-80 countries enrolled in each course, which provided an opportunity to foster intercultural interactions.

2.2 Surveys and interviews

We used a mixed methods research approach to examine MOOC participants' global engaged learning. First, we conducted the Global Engagement Survey with all course participants, including scales that measure intercultural competence (communication and self-awareness) and civic engagement (efficacy, political voice, values and conscious consumption). Second, we conducted semi-structured interviews with a selected sample of course participants to gain a deeper understanding of participants' global learning experience and perpectives.

3 IMPLICATIONS / FUTURE WORK

The results of the proposed research will provide insights about measuring MOOC participants' non-cognitive learning outcomes, and offer suggestions for designing and teaching MOOCs to foster participants' global engaged learning.

REFERENCES

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). *Engaging with massive online courses.* Paper presented at the Proceedings of the 23rd international conference on World wide web.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, New Jersey: Prentice Hall.

Braskamp, L. A., Braskamp, D. C., Merrill, K. C., & Engberg, M. (2008). Global Perspective Inventory (GPI): Its purpose, construction, potential uses, and psychometric characteristics. *Retrieved January, 12*, 2010.

Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE transactions on Learning Technologies, 7*(4), 346-359.

Hartman, E., & Kiely, R. (2014). Pushing boundaries: Introduction to the global service-learning special section. *Michigan Journal of Community Service Learning, 21*(1), 55-63.

Hollands, F. M., & Tirthali, D. (2014). *MOOCs: Expectations and Reality. Full report*. Retrieved from New York: https://files.eric.ed.gov/fulltext/ED547237.pdf

Kellogg, S. B., Booth, S., & Oliver, K. (2014). A social network perspective on peer supported learning in MOOCs for educators. *The International Review of Research in Open and Distributed Learning, 15*(5), 263–289.

Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science, 355*(6322), 251-252.

Liu, Z., Brown, R., Lynch, C. F., Barnes, T., Baker, R., Bergner, Y., & McNamara, D. *MOOC learner behaviors by country and culture; An exploratory analysis*. Paper presented at the Proceedings of the 9th International Conference on Educational Data Mining.

Lough, B. J., McBride, A. M., & Sherraden, M. (2009). Measuring volunteer outcomes: Development of the international volunteer impacts survey. *Center for Social Development, Washington University, St Louis, MO*, 09-31.

Morais, D. B., & Ogden, A. C. (2011). Initial development and validation of the global citizenship scale. *Journal of studies in international education, 15*(5), 445-466.

Paige, R. M., Fry, G. W., Stallman, E. M., Josić, J., & Jon, J. E. (2009). Study abroad for global engagement: the long - term impact of mobility experiences. *Intercultural Education, 20*(sup1), S29-S44. doi:10.1080/14675980903370847

Reynolds, N. P., & Hartman, E. (2016). *Global Engagement Survey*. Retrieved from http://globalsl.org/ges/

Wals, A. E. J. (2007). *Social learning towards a sustainable world: Principles, perspectives, and praxis*. Wageningen, The Netherlands: Wageningen Academic Publishers.

Involving Stakeholders in Learning Analytics: Opportunity or Threat for Learning Analytics at Scale?

Tinne De Laet
Faculty of Engineering Science, KU Leuven
tinne.delaet@kuleuven.be

Tom Broos
Faculty of Engineering Science, KU Leuven
tom.broos@kuleuven.be

Katrien Verbert
Department of Computer Science, KU Leuven
tom.broos@kuleuven.be

Jan-Paul van Staalduinen
Delft University of Technology
J.P.vanStaalduinen@tudelft.nl

Martin Ebner
TU Graz
martin.ebner@tugraz.at

Philipp Leitner
TU Graz
philipp.leitner@tugraz.at

ABSTRACT: This article introduces the goal and activities of the **LAK 2018 half-day workshop** on the involvement of stakeholders for achieving learning analytics at scale . The goal of the **half-day workshop** is to gather different stakeholders to discuss at-scale learning analytics interventions. In particular the workshop focuses on learning analytics applications and learning dashboards that go beyond the implementation in a single course or context, but that have at least the potential for scaling across different courses, programs, and institutes. The main theme of the workshop is to explore how the involvement of different stakeholders can strengthen or hinder learning analytics at scale. The key findings, recommendations, and conclusions of the workshop will be presented in a summarizing report, which will be shaped as a SWOT analysis for stakeholder involvement for achieving learning analytics at scale.

Keywords: Scalability; Institutional implementation; adoption, learning analytics, stakeholder involvement

1 THEME AND WORKSHOP BACKGROUND

Learning Analytics (LA) is relatively young discipline that has gathered promising results. However, these promising results have not yet resulted in widespread implementation in practice. Often learning analytics tools have difficulty to move out of their prototype setting into the real educational practice. It has proven to be challenging to create scalable implementations of learning analytics in authentic contexts that go beyond a particular course or setting (Ferguson et al., 2014). Ethics, privacy (Pardo & Siemens, 2014), technical implementation, integration with existing systems, etc. introduce hurdles for implementation in practice and at scale (Khalil, Khalil, & Ebner, 2015). By involving the different institutional stakeholders in the development, testing, deployment, and assessment phase of learning analytics tools, these hurdles might already be mitigated in an early stage of the project. This workshop aims at collecting experiences of implementing learning analytics applications and learning dashboards at scale and the explicit role of different stakeholders in this process (Drachsler & Greller, 2012). The workshop collects “best practices” and “points for improvement” from the diverse LAK community so that the findings can be shared within the community to boost the future implementation of learning analytics at scale.

In the workshop we would like emphasize three viewpoints:

- **Actual experience viewpoint:** experiences with real-life case studies of learning analytics applications or dashboards who actually have been deployed, or have the potential to be deployed at large scale.
- **Technology-wise viewpoint:** technology for learning analytics at scale and integration with existing (proprietary?) school or higher education systems
- **Stakeholder involvement:** how to involve stakeholders for building an institutional or national policy that can pave the road for learning analytics at scale.

For any of the above themes the contributions of researchers as well as practitioners are welcomed. To facilitate comparison and generalization, all submissions will have to be organized according to the recommendation of Bodily and Verbert (Bodily & Verbert, 2017), who recommend nine categories for describing student-facing learning analytics dashboards, and the general framework of learning analytics of Greller and Draschler (Greller & Drachsler, 2012), who use six critical dimensions to describe learning analytics. An example paper will be provided to assist the authors in adhering to these guidelines.

Submissions with actual evaluations results are stimulated, especially if they use state of the art learning analytics evaluation frameworks, such as the one proposed by Scheffel (“Evaluation Framework for LA - LACE - Learning Analytics Community Exchange,” n.d.; Scheffel, 2017).

2 WORKSHOP DETAILS

2.1 Type of event

The half-day workshop includes different activating formats. The different types of activities in the workshop are focused on achieving one final goal: a SWOT analysis for stakeholder involvement for achieving learning analytics at scale.

The workshop will use an innovative format to ensure that all participants are well-prepared and will be active before and during the workshop. First, rather than presenting their own work, attendees will be asked to present the work, using a presentation or a poster/handout, of another participants based on the publication that was submitted to the workshop. Secondly, another participant (the 'discussant') will be asked to prepare three questions about the submission to be send to all other participants. These questions will also be published on the workshop's webpage.

2.2 Type of participation and target group

The workshop aims at a wide target group: practitioners, policy makers, student representatives, researchers, educational managers from higher education, etc.

The workshop welcomes **two kinds of participants**: contributors with a presentation or poster and contributors interested in sharing their experiences and joining the discussion.

3 OBJECTIVES & PLANNED OUTCOMES

All the presentations and posters from the workshop will be published on the workshop project page, hosted on the project webpage (Erasmus+ project STELA http://stela-project.eu/LAK2018-workshop). The discussion at the workshop will be documented and published on the webpage. Most importantly the findings, recommendations, and conclusions of the workshop will be presented in a summarizing report. We aim at shaping this as a SWOT analysis for stakeholder involvement for achieving learning analytics at scale. This SWOT analysis will be an integral part of the project's outcomes and will be promoted as such. It will be available under open access through the project's and workshop's webpage. The workshop chairs will ensure that the papers presented in the workshop are published in the Companion Proceedings.

4 INTRODUCTION TO ACCEPTED PAPERS

Four papers are accepted for the workshop. Interestingly all papers are a result of ongoing projects ranging from institution-wide projects to European collaboration projects.

4.1 Implementation of an institution-wide learning analytics dashboard: a case study, Ed Foster & Rebecca Edwards

In this paper the authors elaborate on the implementation of a learning analytics dashboard at the scale of an institute: the Nottingham Trent University, United Kingdom. Based on their experiences, they stress the importance of a wide range of stakeholders. analytics tool would not have been possible without the involvement of a broad range of stakeholders. On the positive side, the stakeholder involvement has provided the necessary skills and expertise, triggered new ideas, but also has been key in gaining buy-in and the embedding of the learning dashboard into actual institutional practices. On the negative side, stakeholder involvement has been proven to be time-consuming, has increased the likelihood for miscommunications, and the risk of alienating stakeholders if they feel their feedback is not incorporated.

4.2 Report on a National Learning Analytics Initiative in Ireland; Lee O'Farrell & Geraldine Gray

In this paper the authors elaborate on an ongoing national learning analytics initiative in Ireland. The project fosters collaboration between different higher education institutions, hereby paving the road for campus-wide learning analytics initiatives. The project is using a two stage approach, of which stage one is already complete. The result of the first stage is an online learning analytics information resource for the higher education institutions involved. By involving stakeholders in four working groups, the collaboration across institutions was strengthened and the first steps towards a national learning analytics profile were made.

4.3 Lessons Learned when transferring Learning Analytics Interventions across Institutions; Philipp Leitner, Tom Broos, and Martin Ebner

In this paper the authors elaborate on experiences of transferring learning analytics interventions across institutions within the context of a European project. They focus on a particular case study: the transfer of the learning tracker as developed by project partner TU Delft to KU Leuven, based on a technology stack developed by TU Graz. The challenges are grouped in: working with external providers, and working across institutions. They experiences are grouped in a summarizing table that includes questions that will support future learning analytics transfers to better handle the surfacing challenges.

4.4 The LALA Project: Building Capacity to Use Learning Analytics to Improve Higher Education in Latin America; Jorge Maldonado-Mahauad, Isabel Hilliger, Mar Pérez-Sanagustín, Martijn Millecamp, Katrien Verbert, Xavier Ochoa

In this final paper, interestingly the goals, envisioned approach, and first steps of a new European project LALA ("Learning Analytics Latin America," n.d.) are elaborated on. This project has an even more ambitious goal to transfer learning analytics capacity not just from one European institute to the other, but even from Europe to Latin America. One of the main project goals is to build local capacity and to transfer two specific learning analytics initiatives (LISSA(Charleer, Vande Moere, Klerkx, Verbert, & De Laet, 2017) of the European project ABLE ("ABLE project," n.d.) and the student-facing learning dashboards of the European project ("STELA project," 2017) from Europe to Latin America.

ACKNOWLEDGMENTS

This research project is co-funded by the European Commission Erasmus+ program, in the context of the project 562167EPP-1-2015-1-BE-EPPKA3- PI-FORWARD. The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Please visit our website http://stela-project.eu .

REFERENCES

ABLE project. (n.d.).

Bodily, R., & Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17* (pp. 309–318). New York, New York, USA: ACM Press. https://doi.org/10.1145/3027385.3027403

Charleer, S., Vande Moere, A., Klerkx, J., Verbert, K., & De Laet, T. (2017). Learning Analytics Dashboards to Support Adviser-Student Dialogue. *IEEE Transactions on Learning Technologies*, *10*(3). https://doi.org/10.1109/TLT.2017.2720670

Drachsler, H., & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders, (May), 120. https://doi.org/10.1145/2330601.2330634

Evaluation Framework for LA - LACE - Learning Analytics Community Exchange. (n.d.). Retrieved October 2, 2017, from http://www.laceproject.eu/evaluation-framework-for-la/

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). Setting learning analytics in context. In *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14* (pp. 251–253). New York, New York, USA: ACM Press. https://doi.org/10.1145/2567574.2567592

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, *15*(3), 42–57. Retrieved from http://www.ifets.info/journals/15_3/4.pdf

Khalil, M., Khalil, M., & Ebner, M. (2015). Learning Analytics: Principles and Constraints. *EdMedia: World Conference on Educational Media and Technology*, *2015*(1), 1789–1799. Retrieved from https://www.learntechlib.org/p/151455/

Learning Analytics Latin America. (n.d.). Retrieved February 5, 2018, from https://www.lalaproject.org/

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438–450. https://doi.org/10.1111/bjet.12152

Scheffel, M. (2017). The Evaluation Framework for Learning Analytics. Retrieved from http://dspace.ou.nl/handle/1820/8259

STELA project. (2017). Retrieved from http://stela-project.eu/

Implementation of an institution-wide learning analytics dashboard: a case study

Ed Foster
Nottingham Trent University
ed.foster@ntu.ac.uk

Rebecca Edwards
Nottingham Trent University
rebecca.edwards@ntu.ac.uk

ABSTRACT: The following article provides a case study example of a learning analytics dashboard that has been implemented in a large university in the United Kingdom (~28,000 students). Deployment of the dashboard occurred relatively quickly; from the initial business mandate in early 2013, to a pilot in a subsection of the University in the 2013-14 academic year, to entire institution roll out in the 2014-15 academic year. Efforts from this point have been focused on embedding the tool into institutional culture, and further developing the tool in-line with the needs of the users and the changing environment of Higher Education. The University worked in partnership with the technology provider Solutionpath during the initial design of the dashboard, and continues to collaborate closely. As the institution was an early adopter of learning analytics in the UK, many of the lessons learned have come from direct experience. The article is written from the perspective of the internal business owners of the NTU Student Dashboard, and aims to provide an explanation of the rationale for adopting learning analytics at scale, briefly introduce the reader to the resource, and present an overview of stakeholder involvement throughout the process.

Keywords: Learning analytics, dashboard, case study

1 RATIONALE FOR ADOPTING LEARNING ANALYTICS

Nottingham Trent University (NTU) is one of the largest universities in the UK with approximately 28,000 students studying a range of undergraduate and postgraduate courses in a wide range of disciplines[1]. The majority of students (~75 %) study full time, undergraduate courses. The University has a strong focus on employability and recruits students from a range of socio-economically diverse backgrounds. Undergraduate students enter NTU via two routes; the University & Colleges Admissions Scheme (UCAS), and a process known as 'Clearing'. Like other providers, NTU sets tariffs for its courses; students are accepted based on either fulfillment of the requirements or via negotiations during Clearing. NTU is

[1] https://www.hesa.ac.uk/files/sfr-files/student_sfr242_1516_table_3.xlsx, accessed 26-01-0218

ranked 52nd of 129 UK Universities in the Complete University Rankings 2018[2]. Recently, it has been awarded two nationally-recognised awards; University of the Year at the 2017 Times Higher Education awards and the Modern University of the Year by the Sunday Times[3].

NTU was able to quickly adopt learning analytics as the required data had already been gathered into a data warehouse. Beyond practicalities, the potential of learning analytics to enable a data-driven approach to personalised learning aligned with the University's commitment to strive towards success for students from a diverse range of backgrounds.

In January 2013, following the completion of a major piece of research into student retention (Foster et al., 2012), a business mandate was written to develop a product to achieve the following:

a. Identify students most at risk of withdrawing early and/or underperforming.

b. Trigger a real-time alert for a personal tutor (or other appropriate staff member) when students exhibit at-risk behaviours.

c. Provide clear routes for personal tutors to refer students for additional support.

d. Provide case work capacity to record interventions and manage ongoing student support.

With such a tool, it was envisaged that staff could contact potentially at-risk students earlier to offer support before problems became more serious. On a broader scale, it would allow the institution to research the most effective strategies for interventions and to make informed changes to the curriculum, learning, and teaching to maximise student retention.

In February 2013, the team added a question to the annual Student Transition Survey, which is circulated to all first-year students in the institution. At that early stage, phrasing the question was difficult as it was not clear that students would understand the phrase 'learning analytics' or its implications. Students were asked if the University was able to warn them that they were at risk of early departure would they want to know? 92 % of students (n=441) wanted to be told. This was felt to be sufficient endorsement to proceed.

2 INTRODUCTION TO THE NTU STUDENT DASHBOARD

The NTU Student Dashboard has been shaped significantly by user input at every stage of development from pilot to current iteration. It is a staff- and student-facing resource that provides a view of the

---

[2] https://www.thecompleteuniversityguide.co.uk/league-tables/rankings) accessed 26-01-0218

3 http://www.nottinghampost.com/news/nottingham-news/nottingham-trent-university-named-uks-860414

student's experience at University based on the digital information available. The dashboard presents two forms of data to the user. Firstly, an engagement rating generated by the dashboard's underlying algorithm, analysing students' engagement with learning and teaching activity (for example borrowing a library book). The original version used four data sources and presented four engagement ratings to the user ('high' to 'low'). Following user consultation and the availability of further data, the current version now uses six data sources and presents five ratings ('high' to 'very low'). The data is presented in a way that allows users to see both a summary of engagement behaviour over extended timeframes and short-term changes in behavior (figure 1). Secondly, the dashboard provides valuable contextual information including a profile page detailing basic information about a student including a photograph, their entry qualifications, and their course description and personal tutor's name. Other pages provide assessment and feedback documentation for coursework submitted through the VLE, and the ability to make/view notes about staff-student meetings and subsequent actions or referrals to support.

The dashboard works on a two agent of change model; students can use the data for benchmarking, self-reflection and goal setting, and staff can use the data to assist students in these processes and to identify students who require additional support. This is perceived as particularly useful for first year students who are new to Higher Education. It should be noted that whilst any approach using purely digital data is likely to have limitations, the tool can still be used effectively within these. At NTU, one of the key functions of the dashboard is to strengthen the staff-student relationship by facilitating informed discussions. As a student facing tool, the language contained within the dashboard is very important. In the pre-project phase it was decided that the dashboard should focus on student engagement, not risk of failure. It was felt that if a student could see on their dashboard that they were highly at risk of withdrawal, this could be potentially highly demotivating and counterproductive. Therefore, 'high' means 'highly engaged' not 'highly at risk'.

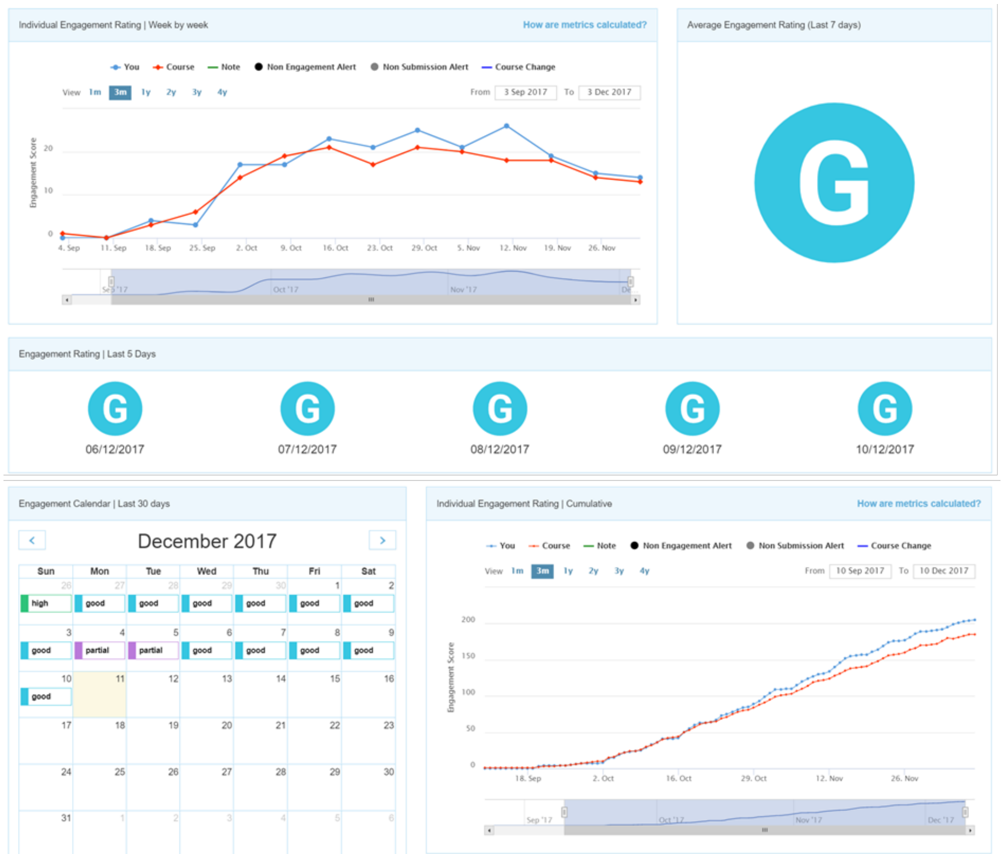


Figure 1: Screenshot of the NTU Student Dashboard student landing screen, showing different visualisations of the engagement data.

3 STAKEHOLDER INVOLVEMENT

As an institution-wide resource, the dashboard has a wide range of internal stakeholders. Arguably, most key are the users themselves: students and staff. Staff members include course staff, particularly personal tutors, and professional services staff members providing pastoral and academic support. However, the interests of the users need evaluating and prioritising. This work is carried out by the NTU business owners; the Student Engagement Team, and the University's Information Systems (IS), who are jointly responsible for the design, implementation and maintenance of the tool. Strategic vision, provision of resource, and high-level backing is provided by University senior managers, and is vital to the continued success of the project.

Further stakeholders include the data owners of the information used by the tool, the policy makers who are responsible for outlining expectations of use of the tool, and the legal services team who are responsible for ensuring the tool and its use comply with legislation. The providers of the dashboard, Solutionpath, are the major external stakeholders in the process, as developing and maintaining a high-quality product is important to their business strategy.

The structures and mechanisms in place to facilitate stakeholder involvement have proved invaluable throughout the process of dashboard development and integration into institutional practice. A combination of different types of working groups provided the necessary access to a broad range of staff members, including specialist support from relevant areas of the University and Student Union representatives, at the point where they could most usefully contribute to the resource's development (Figure 2). Further views, invaluable to the development process, were elicited from staff and student users by the central market research team. Further to these formal structures, stakeholder involvement has been promoted by the running of small to medium scale pilot activities with the academic Schools and gathering feedback from user queries to the IS service desk.

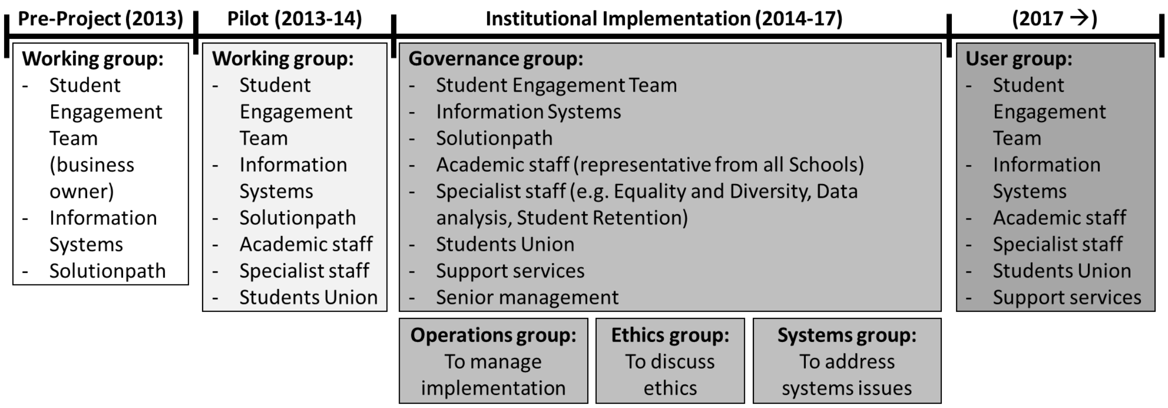


Figure 2: Working groups in place during the various stages of dashboard implementation

The core principles that underpin the dashboard were defined by the pre-project team in consultation with staff, students and relevant specialists from within the institution, and have remained the same since. These well-considered and ethically-informed decisions were perceived as strong foundations upon which the tool should be built. The core functionality, in terms of building the algorithm itself and the timescale for appropriate alerting, was a data-driven process led by Solutionpath and, with the exception of a recent change of the algorithm to incorporate additional data sources, has also remained consistent. This reflects the need to focus on expert input for certain aspects of a project, and to have a framework to underpin supporting features and future developments.

Perceptions of the dashboard are important to stakeholder engagement and buy-in. Designing the product with both the core purpose and the ethical issues at the forefront, and with a high degree of user input, can help mitigate against concerns. User views have shaped the resource at every stage of dashboard development. Many integral elements of the tool, including the information it displays, the way it presents data, the language used within it, and its additional functionalities, have been guided by user input. Developing the tool in this way has undoubtedly led to a more useful resource, however, it has not been problem-free. It is far easier to gather user views than it is to actually integrate them into an existing schema, let alone provide development capacity to fulfil them. Gathering user views without the capacity to quickly embed them was at times frustrating and demotivating to the development team. Secondly, users have blind spots. Staff in particular, were fixated on the importance of attendance data.

Despite repeated evidence of the association between engagement and student success (without attendance data), there remained in some quarters the view that attendance was the more important measure and that the algorithm could not function correctly without it. This may reflect communication failings by the development team, but also demonstrates how difficult it could be to encourage people to think differently.

Early feedback from students highlighted the need to scaffold the introduction of the learning analytics tool in different ways. Overall, students liked the resource, but it was clear that they had not particularly engaged with it during the pilot. In part this was because the link to access the tool, although not hidden, was not obvious enough. Students wanted more communication about how to use the resource and the benefits of doing so. They wanted to see the same information as staff, for example *"I want to be able to see that is collected on me so I know what they can see about me"*. They particularly wanted to see their own attendance data and finally, they wanted to be able to see the dashboard on mobile devices.

Responding to the needs of the users and committing to developing and promoting the tool and associated resources throughout the implementation of the dashboard, has resulted in the number of log-ins increasing year-on-year. From first year of institutional roll out (2014-15) to the last full academic year at the time of writing (2016-17) both the number of unique staff and students users and the average number of log-ins per year has increased. In 2016-17, over 2,500 staff members logged in an average of 16 times per year and over 28,000 students logged in an average of 19 times, taking the total number of log-ins for year to around 600,000. Importantly, in this year, over 90 % of first year, full time undergraduate students across the institution logged in to the dashboard, with over 40 % of these logging in 10 times or more throughout the year.

Involving stakeholders effectively requires a shared language and a shared understanding of both the high-level vision and purpose, and the finer details of the project. Being the business owners at the interface between stakeholders such as IS and Solutionpath; who require high levels of detail to function correctly and stakeholders such as the users and wider University community; who are generally consulted on a much broader basis, has presented its challenges. In particular, being unaware of underlying assumptions and tacit knowledge has led to instances of miscommunication and inefficiencies in the process. One example, from near that start of the development process, was defining which staff members should have access to which students on an institutional level. Whilst University organisational structures and job roles, such as 'personal tutor', may appear clean and easily defined to an outside party, the reality is that they can be both complex and variable across the institution. Courses with non-standard start dates, joint honours degrees and staff with multiple roles are all factors that can make systematically defining which staff should have access to which students at any point in time more challenging than it may first appear, particularly if the system is not built with these complexities in mind.

4 CONCLUSION

Successful implementation and integrating of an institution wide learning analytics tool would not have been possible without the involvement of a broad range of stakeholders. Involving stakeholders in the process has provided necessary skills and expertise, allowed for new ideas and sense checking, and has been an important part of gaining buy-in and embedding the resource into institutional culture/ working practices. However, it has also been time consuming, has increased the likelihood of miscommunications, and has come with the fear of alienating people when not all advice and feedback could be reflected in the final product.

Experience and common sense both dictate that developing an institutional resource without user views, in particular, would be foolish. However, the importance of experts considering, filtering and developing those users' ideas should not be underestimated. As with any resource, the wider context in which it is being used should be kept in mind, and a balanced and pragmatic approach to should be taken to its development.

ACKNOWLEDGMENTS

This research project is co-funded by the European Commission Erasmus+ program, in the context of the project 2015-1-UK01-KA203-013767. The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. Please visit our website http://www.ableproject.eu/.

REFERENCES

Foster, E., Lawther, S., Lefever, R., Bates, N., Keenan, C. and Colley, B. (2012). HERE to stay? An exploration of student doubting, engagement and retention in three UK universities. In: I. Solomonides, A. Reid and P. Petocz, eds., Engaging with Learning in Higher Education. UK: Libri Publishing, pp. 95-114.

Report on a National Learning Analytics Initiative in Ireland

Lee O'Farrell
National Forum for the Enhancement of Teaching and Learning in Higher Education
Lee.OFarrell@teachingandlearning.ie

Geraldine Gray
Institute of Technology Blanchardstown
Geraldine.Gray@itb.ie

ABSTRACT: Development of learning analytics capacity and practice at institution level is a challenging task. This paper reports on an ongoing, national project in Ireland that is addressing this challenge by fostering learning analytics collaborations between higher level institutions. Such collaborations are enabling the development of learning analytics capacity across the higher education sector with a common goal of supporting a holistic view of student success. Academic and non-academic staff from over twenty Higher Education Institutions are involved in the project.

Keywords: institutional approach to learning analytics, student success initiatives, higher education in Ireland.

1 INTRODUCTION

Cross-institution coordination of learning analytics is rare (Shacklock, 2016), and can be a daunting task from a number of perspectives including financial costs, ethical and privacy considerations, and uncertainty on beneficial uses of student data and models (Ferguson et. al, 2016; Slade & Prinsloo 2013). A recent review of Learning Analytics in the thirty-six Higher Education Institutions (HEIs) in Ireland reflects experiences elsewhere (O'Farrell, 2017). To some extent, all HEIs were using learner data to understand and respond to students' learning needs. Simpler examples included identification of at-risk students from assessment data at subject level; use of grade curves at program level to identify modules that deviated from the normal distribution; and reviews of student services based on qualitative data generated through feedback surveys. However, just three Irish HEIs were using learning analytics as part of an institutional strategy focused on student success and retention; three had mature learning analytics capabilities that were not driven strategically at institution level; five were planning technologies or data management approaches to enhance learning analytics capability; and just one staff member held a role that formally included intervening with students whose digital footprint suggested a lack of engagement. The thirty-six HEIs included seven publically funded universities, fourteen publically funded Institutes of Technology, six partly funded colleges and nine private, non-for-profit colleges.[1] Interestingly, strategic institutional

[1] Applications to all full-time higher education courses in Ireland is managed by a Central Applications Office. Places are offered based on grades achieved in state examinations at the end of secondary school. Fees are comparable across the sector, and capped at €3,000 per annum for EU citizens.

learning analytics approaches were more common in private HEIs than publicly funded institutions, which reflects evidence from other jurisdictions (Sclater & Mullan, 2017).

Commonly cited barriers to greater adoption of learning analytics in Irish institutions included: limited resources with other business critical priorities taking precedence; a perception of lack of expertise in developing learning analytics capacity; a perception that learning analytics requires a significant capital investment; and a lack of awareness of learning analytics capacity within VLE platforms already in use. However, learning analytics was ranked within the top five institutional priorities for Irish HEIs over the next three years (National Forum, 2017).

This paper reports on a national project established to collaboratively develop learning analytics capacity in the HEI sector in Ireland. The project is led by the National Forum for the Enhancement of Teaching and Learning in Higher Education[2], a government funded agency tasked with enhancing teaching and learning for all students in higher education. The paper also includes a perspective from one HEI involved in the project, Institute of Technology Blanchardstown.

2 NATIONAL LEARNING ANALYTICS INITIATIVE

Overarching the developing of a national led learning analytics initiative is adherence to principles of moral and ethical standards of practice as proposed by Slade & Prinsloo (2013). These include: recognition that students are active partners in the learning process; full disclosure of all uses of data (a legal requirement in the EU); realization that analysis of learning data cannot accurately reflect the complexities of learning behavior and so is limited in scope; and recognition that analysis of data provides a snapshot of some aspect(s) of student behavior that is itself fluid, so is limited temporally. Therefore, it is important to differentiate between a reporting of facts, and understanding that learning is more complex than the facts and models can capture. Without a more nuanced, considered understanding, the risk of commodifying students becomes a genuine threat to how students are perceived and treated by the institution. To paraphrase Kahneman (2011), “what you see is *not* all there is”.

In spite of these limitations, learning analytics has gathered considerable momentum, and its potential for enhancing teaching and learning has been correctly lauded. Learning analytics models can provide useful insights into the learning environment, and can support teaching and learning if used in combination with effective intervention strategies (Dawson, Jovanovic, Gašević, & Pardo, 2017; Jayaprakash, Moody, Lauría, Regan, & Baron, 2014). Analytics is a tool to help answer questions and provide insights. Effective use of this tool in broader, proactive student success initiatives can enhance the learning experience.

As the principles outlined above emphasize, institutions will have a better chance of achieving broad student success if their learning analytics strategies consider the whole student, the dynamic nature of learning, and the conditions within which learning occurs. Individual differences dictate that one-size-fits-

[2] https://www.teachingandlearning.ie/

all interventions risk being insensitive to the entire range of extra-academic issues that individual students face. Such intervention models may also promote a false understanding that all students, regardless of personal circumstances, must adhere to a given model of learning and success (O'Farrell, 2017).

Within this context, in 2016 the National Forum for the Enhancement of Teaching and Learning in Higher Education launched its Learning Analytics and Educational Data Mining for Learning Impact project[3], led by Lee O'Farrell. The project aims are:

- To raise awareness of emerging national and international policy and practice relating to Learning Analytics / Education Data Mining (LA/EDM) in all sectors of higher education in Ireland, among the student body, library/learning support staff, ICT/services staff and academic staff at all levels engaged in developing and teaching programmes and in senior/academic leadership roles.
- To establish a sustainable network of LA/EDM practitioners/collaborators in Irish HE with a view to proactive information sharing and development and dissemination of relevant case studies.
- To provide informative briefings that can support the translation of LA/EDM research findings (national & international) into meaningful practice at scale within programmes/departments in different academic disciplines.
- To develop a set of online resources (including links to already existing resources) relating to LA/EDM policy, practice and implementation at scale.
- To foster intra- and inter-institutional collaboration in the development and implementation of LA/EDM initiatives, with particular reference in the first instance to implementations that target first year student retention.

The project has two phases: phase one developed an Online Resource for Learning Analytics (ORLA); phase two is working with HEIs on a Data-Enabled Student Success Initiative (DESSI). Both will be discussed in the following sections, along with the experiences of one partner HEI.

2.1 ORLA: Online Resource for Learning Analytics

The Online Resource for Learning Analytics (ORLA)[4] incorporates a range of learning analytics resources relevant to higher education, and is summarized in Figure 1. Launched in October 2017, ORLA includes guidelines on how to develop an institutional learning analytics strategy, how-to guides for educators, and learning analytics case studies from Ireland and abroad. The resulting resources arose from the work of four national advisory groups convened to contribute to ORLA, comprising of sixty representatives from eighteen HEIs across Ireland. The groups were: *IT & Infrastructure* to document the data captured by

[3]Project website: https://www.teachingandlearning.ie/priority-themes/mapping-digital-capacity/pre-specified-nationally-coordinated-projects/learning-analytics-and-educational-data-mining-for-learning-impact/

[4] OLRA website: www.teachingandlearning.ie/NFOrla

platforms and products in use across the HEI sector in Ireland including student information systems, library systems, VLEs, and other data sources; *Data Aggregation & Modelling* to develop guides on modelling this data; *Policy, Ethics & Law* to develop guidelines on ethical and legal considerations; and *Teaching, Learning and Effective Interventions* to develop guidelines on good practice when planning analytics led interventions.

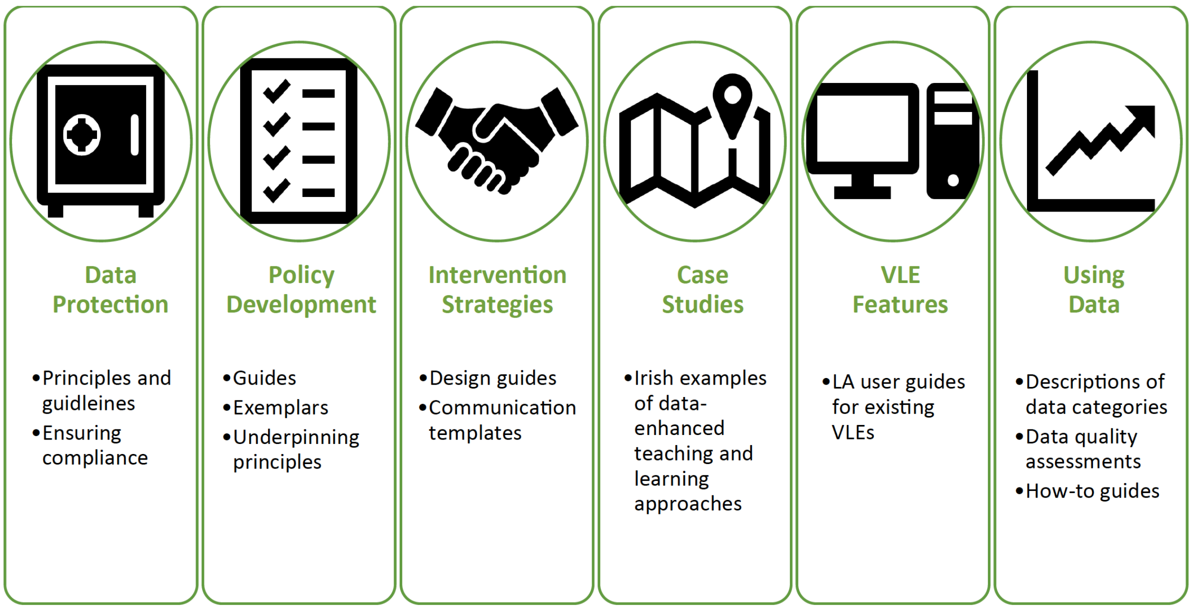


Figure 1. Online Resource for Learning Analytics

2.2 DESSI: Data-Enabled Student Success Initiative

The Data-Enabled Student Success Initiative (DESSI)[5] will progress the work of ORLA by working with individual institutions to develop their learning analytics capacity. In line with the national imperative to pool resources and share services (Department of Education and Skills, 2017), the project will identify common requirements across institutions that can be developed at a national level, including policy recommendations and/or tools. It is led by the National Forum for the Enhancement of Teaching and Learning in Higher Education, in partnership with the Department of Education and Skills and state agencies supporting and overseeing Higher Education in Ireland. These agencies include the two overarching bodies of the Higher Education Authority (HEA) and Quality and Qualifications Ireland (QQI); the three bodies overseeing sections within HE, namely, the Irish Universities Association (IUA), the Technological Higher Education Association (THEA), and the Higher Education Colleges Association (HECA); the two bodies providing IT services to the education sector, namely, HEAnet (providing internet connectivity and ICT services) and EduCampus (provider of IT shared services); and the Irish Survey of Student Engagement (ISSE) group who run national surveys for student feedback. The project is funded until the end of 2018, and is guided by four core principles: developments in learning analytics should

[5] www.teachingandlearning.ie/DESSI

support a holistic view of student success; taking a strategic institutional approach to learning analytics is both valuable and necessary; resources, tools and services should only be employed by HEIs to support learning analytics following a review of their suitability, scalability and adaptability to the specific context; and every effort should be made to share learning analytics services across the sector to avoid inefficiencies or duplication of effort.

It is intended that the spirit of collaboration fostered in the development of ORLA will be harnessed within DESSI to allow institutions to efficiently foster innovative, evidence-based teaching and learning environments with student success at their core.

2.3 The perspective from one HEI: Institute of Technology Blanchardstown

Those involved directly in ORLA and DESSI comprise of a small number of staff from any one HEI. However, successful implementation of a campus wide learning analytics initiative requires engagement by all relevant stakeholders within an institution. This section case studies activities arising from ORLA and DESSI at one HEI, namely Institute of Technology Blanchardstown (ITB). It is the newest Institute of Technology in Ireland, established in 1999. It offers undergraduate and post graduate courses in Computing, Engineering, Horticulture, Social Care, Business and Sports Management. The college has an enrollment of approximately 3,500 students.

Following the launch of ORLA, ITB ran two workshops to start the conversation on ITB's learning analytics agenda, and collect perspectives from relevant stakeholders across the institution. All staff were invited. A total of twenty-eight attended workshops, including academic staff and representatives from student support, library, exams office, finance office, IT services, quality assurance, careers service, student representatives and senior management. The following paragraphs summarize workshop outcomes, capturing campus wide perspectives and aspirations for learning analytics.

Understanding our students: Learning analytics should enhance our ability to understand students, inform support initiatives, and promote an ethos of humanity and empathy. Data analysis should incorporate the full student story including non-academic contexts such as where they come from, how they get here, are they spending time on campus and are they joining clubs and societies.

Understanding our data: We need a better understanding of the data we have and better understanding of what can be inferred from existing data sources. There is also scope for greater awareness of ways to improve data quality, such as the use of more descriptive names on learning resources to improve the usefulness of data recorded in VLE activity logs.

Enhance student support: Learning analytics should enable early identification and follow up support of at-risk students, and provide more opportunities for student feedback including sentiment.

Limitations: Time and financial constraints dictate small steps with simple analytics as the most realistic next steps.

ITB's involvement in ORLA and DESSI has raised the profile of learning analytics across the institute, and has fostered invaluable momentum, support and enthusiasm for the development of campus-wide learning analytics capacity. Workshop participants are currently establishing a learning analytics committee with representation from student support/library/careers, exams/finance, IT services, quality assurance, student representative and academic staff. It will report to the quality assurance subcommittee of academic council. The learning analytics committee will consider both workshop outcomes, and ITB student expectations and perceptions currently being assessed using the Student Expectations of Learning Analytics Questionnaire (SELAQ) developed by Whitelock-Wainwright, Gašević and Tejeiro (2017). Working with other partners in DESSI, a priority of this group is to develop a learning analytics policy and strategy, student information sheet, and a data protection impact assessment for learning analytics at ITB. The group will also identify a first campus-wide learning analytics project that supports existing First Year Experience initiatives.

3 CONCLUDING REMARKS

Fostering collaborations across the higher education sector can support the development of campus-wide learning analytics initiatives. In Ireland, a national led Learning Analytics and Educational Data Mining for Learning Impact project is doing this in two stages. The first stage convened four working groups to develop an online, learning analytics information resource for the HE sector. This both fostered collaborations across institutions, and raised the national profile of at-scale learning analytics. Stage two is now underway, working with learning analytics committees at partner institutions to identify common requirements, including both policy recommendations and tools. A case study of one institution highlighted the benefits of both a national agenda, and a collaborative approach, in building momentum and support for the development of institution wide learning analytics capacity.

Cross-institution coordination of learning analytics is a non-trivial task, that requires the co-operation and input from students, academic staff, support staff, and management. Resource limitations support less costly, phased based implementation approaches, informed by guidelines on some simple but effective uses of data and a better understanding of resources already in place. There is a lot of expertise dispersed both within and across institutions. Collaborative initiatives can consolidate existing expertise while maintaining an overarching, common goal to keep a focus on the human stories behind the data, and ensuring that facts and figures alone do not become the full story.

REFERENCES

Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017, March). From prediction to impact: evaluation of a learning analytics retention program. In *Proceedings of the 7th International Learning Analytics & Knowledge Conference* (pp. 474-478). ACM.

Department of Education and Skills. (2017). *Education and training sector shared services plan 2017-20*. Dublin: Department of Education and Skills.

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., & Vuorikari, R. (2016). Research Evidence on the Use of Learning Analytics: Implications for

Education Policy. Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC104031/lfna28294enn.pdf

Shacklock, X. (2016). *From bricks to clicks – The potential of data and analytics in higher education*. Higher Education Commission. Retrieved from http://www.policyconnect.org.uk/hec/research/report-bricks-clicks-potential-data-and-analytics-highereducation

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6-47.

Kahneman, D. (2011). *Thinking, fast and slow*. London: Penguin.

National Forum. (2017). *Ireland's higher education technical infrastructure. A review of current context, with implications for teaching and learning enhancement*. Dublin: National Forum for the Enhancement of Teaching and Learning in Higher Education.

O' Farrell, L. (2017). *Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher Education*. Dublin: National Forum for the Enhancement of Teaching and Learning in Higher Education.

Sclater, N., & Mullan, J. (2017). Learning analytics and student success – assessing the evidence. Retrieved from https://repository.jisc.ac.uk/6560/1/learning-analytics_and_student_success.pdf

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529.

Whitelock-Wainwright, A., Gašević, D., & Tejeiro, R. (2017, March). What do students want?: towards an instrument for students' evaluation of quality of learning analytics services. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 368-372). ACM.

Lessons Learned when transferring Learning Analytics Interventions across Institutions

Philipp Leitner
TU Graz
philipp.leitner@tugraz.at

Tom Broos
KU Leuven
tom.broos@kuleuven.be

Martin Ebner
TU Graz
martin.ebner@tugraz.at

ABSTRACT: Learning Analytics is a promising research field, which is advancing quickly. Therefore, it finally impacts research, practice, policy, and decision making [7] in the field of education. Nonetheless, there are still influencing obstacles when establishing Learning Analytics initiatives on higher education level. Besides the much discussed ethical and moral concerns, there is also the matter of data privacy.

In 2015, the European collaboration project STELA started with the main goal to enhance the **S**uccessful **T**ransition from secondary to higher **E**ducation by means of **L**earning **A**nalytics [1]. Together, the partner universities develop, test, and assess Learning Analytics approaches that focus on providing feedback to students. Some promising approaches are then shared between the partner universities. Therefore, the transferability of the Learning Analytics initiatives is of great significance.

During the duration of our project, we found a variety of difficulties, we had to overcome to transfer one of those Learning Analytics initiatives, the Learning Tracker from one partner to the other. Despite, some of the difficulties can be categorized as small, all of them needed our attention and were time consuming. In this paper, we present the lessons learned while solving these obstacles.

Keywords: Learning Analytics, scalability, cooperation, lessons learned

1 INTRODUCTION

Learning Analytics has emerged in the last decade as a fast-growing and promising research field in Technology-Enhanced Learning (TEL) by providing tools and platforms that influence researchers [10, 6]. Long defined Learning Analytics as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environment in which it occurs” [14]. Since it was first mentioned in the Horizon Report 2012 [9], various different projects and initiatives were performed surrounding Learning Analytics, which is finally entering the next phase and has an impact on research, practice, policy, and decision making [7].

Nonetheless, there are many obstacles when establishing Learning Analytics initiatives especially in higher education. Besides ethical and moral issues, the matter of data ownership and data privacy is getting more and more important [5]. Particularly affected are the member states of the EU as the new EU General Data Protection Regulation (GDPR)[1] is going to be enforced soon. Thereby, the users, lecturers and students, have to be informed in advance of what is going to happen with their personal data as well as give the consent. Unfortunately, anonymizing personal data to circumvent the issue with personal data makes Learning Analytics more difficult and is not that trivial [11]. Further, many Learning Analytics projects are still in the prototype phase, because of issues with transferability and scalability [13].

Within the scope of the European collaboration project STELA, the Learning Tracker [8] was proposed for giving students feedback in a Small Private Online Courses (SPOC) deployed at KU Leuven. In this publication, we will present issues and lessons learned in the process of deployment. We summarized this through two research questions:

RQ1: What should be kept in mind when working with external providers?

RQ2: What should be kept in mind when working across higher education institutions?

In the next section, we start by explaining the case study and its circumstances. Section 3 explores issues when working with external providers and the lessons learned. In Section 4, we discuss obstacles when working across institutions and how to overcome them. Conclusion and remarks on future work are presented in Section 5.

2 CASE STUDY

The Erasmus+ STELA project [1] is a European collaboration project with the primary partners Catholic University of Leuven (KU Leuven, Belgium), Delft University of Technology (TU Delft, Netherlands), Graz University of Technology (TU Graz, Austria), and as secondary partner the Nottingham Trent University (NTU, England). The main goal is to enhance the successful transition from secondary to higher education by means of learning analytics. Together, the partner universities develop, test, and assess Learning

1 https://www.eugdpr.org/ - Last accessed January 30th, 2018

Analytics approaches that focuses on providing formative and summative feedback to students in the transition. In the first step, promising approaches are shared between the partners to evaluate them under different circumstances. Therefore, transferability, scalability, and modularity of the approaches are of high interest.

One promising initiative of University of Technology of Delft is the so-called "Learning Tracker" [8], which is made available by TU Delft as open source[2] and is displayed in Figure 1 by a typical user interface. The Learning Tracker itself tracks the behavior of all current participants in the MOOC and displays it against the aggregated activities of previous participants that successfully completed. Thereby, the Learning Tracker supports learners in Massive Open Online Courses (MOOC) in becoming more efficient and encourages them to develop their self-regulated learning skills by reflecting on their own learning activities [8]. This approach follows Baker's alternate paradigm for online learning by using the information to rather empower human decision making than feeding it to an intelligent learning system [2].

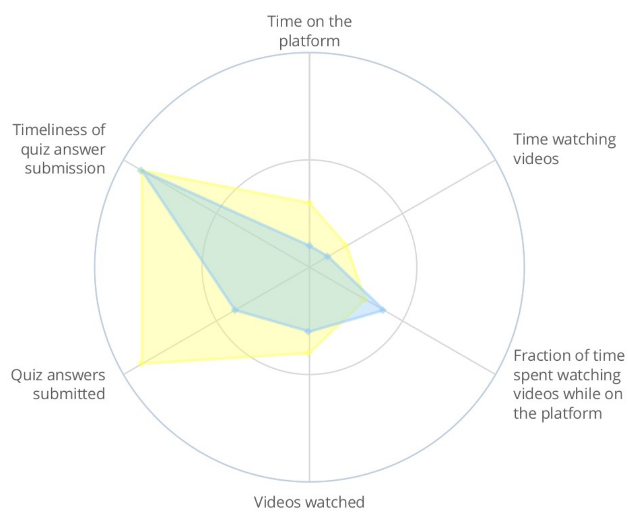


Figure 1: Visual design of the Learning Tracker. It provides several metrics in a small space and offers a simple overall evaluation [8]

The Learning Tracker was already deployed within different MOOCs and has been shown to be easily transferable to different MOOCs on the same platform within the same university [4]. The impact on the engagement of the students in comparison to the completion rate of the MOOC was evaluated and the results have shown that the Learning Tracker improves the achievement of already highly educated learners, but is less effective for less educated ones [4]. Further, it has been shown that the cultural context of the learners is impacting the engagement and the completion rate [4].

2 https://github.com/ioanajivet/LearningTracker – Last accessed January 30th, 2018

Our goal was to deploy the Learning Tracker to the Chemistry SPOC of KU Leuven, which is based on the edX system. Further, we wanted to get the Learning Tracker more dynamically. Therefore, we used the opensource technology stack developed at TU Graz within the STELA project [12]. Figure 2 illustrates a flow diagram of responsibilities and relations throughout the case study.

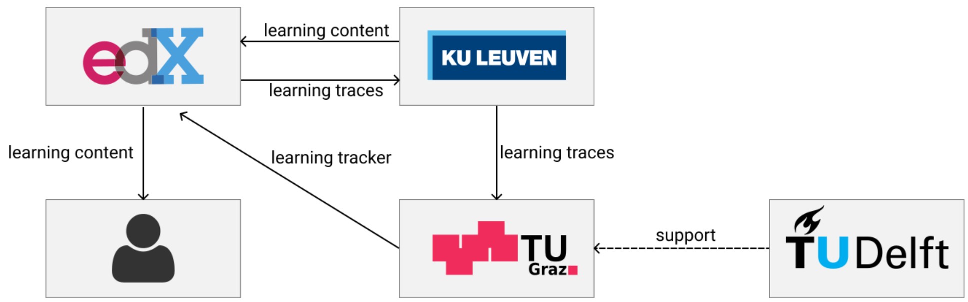


Figure 2: Illustration of responsibilities and relations

3 WORKING WITH EXTERNAL SERVICE PROVIDERS

This section deals with obstacles when working with external service providers (RQ1). We start by explaining issues with data ownership when using an external service provider. Then, we discuss what should be kept in mind when exchanging data with external service providers.

3.1 Data ownership issues

Essential when working with external service providers, is the question ”who owns the data?”. Here we don’t consider matters related to copyright of the material provided on the platform. We also make abstraction of the more fundamental idea that the final ownership of student produced data, whether it concerns learner created content or simply digital activity traces, should always belong to the students themselves.

When the external party functions as a contractor for the institution, it is reasonable to assume that the latter preserves full ownership. But what if the platform of the service provider is independent and subsequently used by the institution to deliver its content to learners?” To draw a parallel: when a company uses a popular social media platform like LinkedIn to disseminate its message, would one not assume that the platform provider retains the ownership of data related to its own user base, even if it was in fact the company that pushed these users to the platform in the first place? And yet, it may come as a surprise to institutions that they don’t automatically acquire ownership of or even access to student data within the external educational platforms used by them.

KU Leuven invested extensively in its Learning Management System ”Toledo”, which is predominantly based on the Blackboard product line. The system is maintained by an internal team and embedded in a broader software architecture, fully hosted in the university’s own data center. Only in recent years, KU Leuven started to invest in MOOCs and SPOCs. Due to the limited in-house experience with MOOC’s and the common practice of hosting shared by many institutions of using an existing platforms, edX was selected as KU Leuven's MOOC platform of choice. However, while the issue of ownership of "Toledo" data did not arise before, it suddenly become relevant in the new context of the edX platform.

3.2 Exchanging data with external providers

Once an agreement with the external service provider is established, the problem of data access arises. For information systems managed by the institution itself, there is usually an option to extend or modify the software to export data required for the Learning Analytics application. In many cases, the data may also be fetched from a database management system directly, by setting up an ETL-process (extract, transform, load) as is common in the domain of Business Intelligence (BI). Institutional IT services are often familiar with these practices, also used to enable reporting on data captured in financial, administrative, Human Resources (HR), and other information systems.

Yet when working with an external service provider, data is not directly accessible by the internal services. As the software is primarily designed to serve multiple tenants, it may not be straightforward to adapt it to meet the data needs of a single institution – especially in the context of an ongoing research project, when requirements are still unstable.

In some cases, the service provider offers a set of application programming interfaces (APIs) to facilitate the communication with on-premises software of the institutions. However, these APIs are likely to be limited to the use-cases anticipated on beforehand, if not by internal planning and priorities. Innovative and experimental use of data, as it is to be expected within a research context, is not always compatible with this approach. The resulting requirement is to dig deeper into the data that is being captured by the external system, if possible by accessing it directly, circumventing the limited scope of the APIs. After all, this would also be a common approach for internal systems, as explained above.

Apart from the requirement to create clarity about the data ownership and sharing, our case study also involves finding a technical process to get the data from the service provider. edX indeed offers an API for accessing student activity data. However, the provided methods are limited to the data perspectives as imagined by the edX developers and incompatible with the requirements of the TU Delft Learning Tracker. On request, edX offered the option to get direct access to extract the underlying log data through an FTP server. The manual way of working is little optimized for continuous, preferably real-time data extraction, but it allows the initiation of the case study implementation. At KU Leuven side, the process of collecting data from edX needs to be further automated. A question is how to anticipate data structure changes on edX side, as the data format is meant for internal use and might be reorganized in the future.

A related issue concerns the reverse flow: once the exported data has been transformed into information that may be offered to students, how can this information be fed back to them? edX supports the Learning Tools Interoperability (LTI) standard created by the IMS Global Learning Consortium. This standard was designed to enable the sharing of tools across different learning systems. In our setup, the edX environment is the LTI Tool Consumer and our Learning Tracker system is the LTI Tool Provider. When the

Learning Tracker is shown to the student, edX (trusted consumer) passes a user identification string, whom makes an extra authentication step on the provider side unnecessary.

4 WORKING ACROSS INSTITUTIONS

In this section, we discuss obstacles when working across higher education institutions and how to overcome them (RQ2). First, we explain what you need to keep in mind when facilitating cross-border European initiatives. Second, we point out how to handle external data subjects.

4.1 Facilitating cross-border European Initiatives

Research cooperation is common among European universities. Students, lecturers and researchers are increasingly roaming from one institution to another, increasing the opportunities for teaming up. But when the research project directly involves the daily practice of the involved institutions, practical incompatibilities may start to surface.

If working together with institutions within a single region may already be complicated, working across (European) borders is unlikely to make matters easier. Despite the unification efforts of the Bologna Process, Higher Education Institutions (HEI) from different European countries operate in dissimilar contexts. Education and general laws, culture, and societal views on the role of education, organization of the institutions, and role of the government are just a few examples of contextual areas that are likely to differ from one country to another. Not in the least because education today is often influenced by local tradition.

While preparing the case study implementation, it became clear that the Austrian view on data privacy is more strict than the Belgian interpretation. Privacy awareness is stronger developed in the Austrian and German culture. Notwithstanding the General Data Protection Regulation (GDPR), which will soon be in effect throughout the entire European Union, the interpretation of what is allowed and what is not turned out to be rather different. The Austrian reading, as translated into TU Graz internal policy, for instance, directs on avoiding the integration of data from separate source systems.

The concept of processing data about the Belgian students on its Austrian servers provoked resistance on the side of TU Graz, as it would put the internal IT department in a challenging position with respect to its policy. Consequently, the alternative of moving the case study implementation to the KU Leuven infrastructure was considered. However, this would require a TU Graz project member to access the server infrastructure of KU Leuven remotely. While there was no objection to this in principle, this turned out to be practically impossible to arrange without an existing employee relationship: the procedure to do so was nonexistent.

4.2 Handling external Data Subjects

The debate about ethics and privacy in Learning Analytics is growing. Skeptics are questioning to what extend providers of education are entitled to study the learning behavior of their students. LA proponents, on the other hand, are arguing that it is the duty of educators to improve learning and that it not using data to do so may be unethical. In most cases, the (implicit) focus of such debate however, is on data institutions collect and process about their *own* students, it is to say, student with which the institution has some kind of formal engagement. It is not uncommon for students to sign a contract at registration that already contains certain agreements about if and how the institution may use learning traces for educational research or to improve its internal educational processes.

However, as is the situation for our case study it is also not uncommon for higher education institutions to interact with prospective students prior to official registration. This complicates matters of privacy and ethics, and in the absence of an agreement, it is less clear what data institutions can use to extend their mission to improve the quality of education to the orienting and transitional process. We therefore prefer to extract as little data as possible (data minimization) to enable the selected features of the Learning Tracker tool. This, for instance, does not require knowledge of the student's name or any other characteristics, besides some kind of user id or pseudonym which is also required to feed the resulting charts back into the SPOC user interface.

The external data subject issue is discussed in detail by [3], applied there to a shared positioning test for engineers students, co-organized by several universities. The proposed solution uses an anonymous feedback code that is provided to students. In this approach, data subjects retain a large part of the data ownership and freely decide to transfer data across service providers or institutions.

5 CONCLUSION

The intention of this paper was to formulate lessons learned, which the authors consider important for future development and implementation of Learning Analytics initiatives. In this paper we have outlined obstacles when working with external providers (RQ1) or across institutions (RQ2), and proposed partial solutions to overcome them. We try to allow that implementer of Learning Analytics initiatives can benefit from this findings, adjust properly and thereby, save time and effort. In Table 1, a summary of questions that surfaced during our case study is provided.

Table 1: Summary of surfacing questions .

Source of issue	Issue	Question
Working with an external provider	Data ownership	• Who owns the data? The institution or the service provider?
	Data access	• How to get data out of the external platform? Are API's available and sufficient? Is direct data access possible? • How to get information back into the systems? How to reach the end-user? Is a standard (e.g. LTI) supported?

Working across institutions	Working cross-border	• How does the educational context differ from one partner to the other? In case of shared legislation, does the interpretation differ? • What procedures are available to host other partner's data or to provide access to a researcher staffed by another partner.
	External data subjects	• To what extend can data from unregistered/prospective students be used to improve education and to feed information back to these students? • If anonymous data is insufficient, is the use of pseudonymization tokens (e.g. feedback codes [3]) an alternative?

ACKNOWLEDGMENTS

This research project is co-funded by the European Commission Erasmus+ program, in the context of the project 562167EPP-1-2015-1-BE-EPPKA3- PI-FORWARD. The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.
Please visit our website http://stela-project.eu .

REFERENCES

[1] Stela project http://stela-project.eu/ - last accessed january 30th, 2018.

[2] Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600-614.

[3] Broos, T., Verbert, K., Langie, G., Van Soom, C., & De Laet, T. (2018, March). Multi-institutional positioning test feedback dashboard for aspiring students: lessons learnt from a case study in Flanders. In *Proceedings of the Eighth International Learning Analytics & Knowledge Conference* (pp. 1-6). ACM.

[4] Davis, D., Jivet, I., Kizilcec, R. F., Chen, G., Hauff, C., & Houben, G. J. (2017, March). Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 454-463). ACM.

[5] Drachsler, H., & Greller, W. (2016, April). Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 89-98). ACM.

[6] Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5-6), 304-317.

[7] Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64-71.

[8] Jivet, I. (2016). The Learning tracker: a learner dashboard that encourages self-regulation in MOOC learners.

[9] Johnson, L., Adams, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2012). The NMC Horizon Report: 2012 Higher Education Edition. The New Media Consortium.

[10] Khalil, M., & Ebner, M. (2015, June). Learning analytics: principles and constraints. In *EdMedia: World Conference on Educational Media and Technology* (pp. 1789-1799). Association for the Advancement of Computing in Education (AACE).

[11] Khalil, M., & Ebner, M. (2016). De-identification in learning analytics. *Journal of Learning Analytics*, *3*(1), 129-138.

[12] Leitner, P., & Ebner, M. (2017, July). Development of a Dashboard for Learning Analytics in Higher Education. In *International Conference on Learning and Collaboration Technologies* (pp. 293-301). Springer, Cham.

[13] Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—a literature review. In *Learning Analytics: Fundaments, Applications, and Trends* (pp. 1-23). Springer, Cham.

[14] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, *46*(5), 30.

The LALA Project: Building Capacity to Use Learning Analytics to Improve Higher Education in Latin America

Jorge Maldonado-Mahauad, Isabel Hilliger, & Mar Pérez-Sanagustín
Department of Computer Science, Pontificia Universidad Católica de Chile, Chile
{mar.perez, ihilliger, jjmaldonado} @uc.cl

Martijn Millecamp & Katrien Verbert
Department of Computer Science, KU Leuven, Leuven, Belgium
{katrien.verbert, martijn.millecamp}@cs.kuleuven.be

Xavier Ochoa
Centro de Tecnologías de la Información, Escuela Superior Politécnica del Litoral, Ecuador
xavier@cti.espol.edu.ec

ABSTRACT: Learning Analytics has been widely developed in European countries. Latin America is also starting to measure and optimize teaching and learning processes through Learning Analytics. However, the existing attempts in this direction are very isolated. Further efforts are needed in order to promote greater sharing of best practices between regions. Along these lines, the present work presents the LALA project with the aim of building capacity to use Learning Analytics to improve higher education in Latin America. At the end of the project we expect to have built local capacity for decision-making by using educational data, a community of practice around learning analytics, and a methodological framework to guide Latin American institutions to adopt tools to support learning.

Keywords: LALA Project, Learning Analytics, Latin America

1 INTRODUCTION

Students in higher education are producing data at increasing pace. Demographic information captured at registration, test results, interaction with the learning management system (LMS) or which books the student loans from the library are just a few examples of the many data traces left behind by students. Most of this data is being stored for administrative purposes only, often to be never used again. However, these data have the potential to improve the decision-making process of several stakeholders of the higher education institute (HEI).

In Europe, multiple researchers from Higher Education Institutes (HEIs) are looking to support the decision-making process with their available data. This support goes further than reporting data with existing tools as data cubes and ad-hoc queries. This support consists in learning analytics tools that not only report on what happened before, but also estimate what could have happened if the current trend continues to optimize an existing process.

The field of Learning Analytics and Academic Analytics has developed considerably during the last year in European HEIs. Through the use of Learning Analytics tools, the management of Universities and Academic programs in Europe has been modernized. However, in the specific case of Latin American (LatAm), there is a lack of local capacity to design and build these specialized tools that could be used to improve HEIs administration. The scarce of financial resources or the inequalities to distribute them across institutions and educational levels prevent the implementation of tools to analyze educational data, so most decisions in the LatAm academic settings are based on budgeting, preconceptions or even feelings. If data is used, it is in the form of database reports that only provide the most basic level of information. Due to the last decade modernization of academic systems in most LatAm HEIs, there is an opportunity to use that large amount of data to drive management towards learning improvement.

Given that situation, in this work we present the LALA project, which seeks to modernize the academic decision making process through building local capacity to create Learning Analytics to support decision makers at different levels in LatAm HEIs. To achieve this purpose, this project is inspired by several projects currently running in Europe to facilitate the adoption of Learning Analytics tools. These European projects will bring the expertise developed in: 1) ABLE project in which research the use of learning analytics to identify students at risks. The aim of this project relied on the use of learning analytics dashboard that has been designed, developed and evaluated to facilitate communication between study advisors and first year students, 2) STELA project in which research the use of learning analytics dashboards to support interventions to students. The main focus of this project is on providing formative and summative feedback to students, and 3) SHEILA project which aim is to develop a policy development framework that will assist higher education institution in adoption of learning analytics. The policy framework is based on the ROMA model that was originally designed to offer translation of scientific evidence to a policy and recently suggested for learning analytics.

The LALA project fit the priorities set for LatAm inside the Erasmus Plus project call for Capacity Building. That is "Improving management and operation of higher educational Institutions". More specifically, this project seeks to build local capacity in LatAm HEIs to design and implement Learning Analytic tools to create and improve "Quality Assurance processes and mechanisms". Building the local capacity to the design, implementation and use of Learning Analytics tools will provide LatAm HEIs with a powerful tool to solve not only one problem, but any problem where data analysis could be used to inform decision-makers. The main impact expected from the project is the creation of local capacity to create, adapt, implement and adopt Learning Analytics tools to improve the academic decision-making process on LatAm HEIs.

To let this project succeed, a community of practice will be built, starting by different HEIs from LatAm and Europe that combine knowledge and experience. Concerning project partners, there are two universities from Chile: Pontificia Universidad Católica de Chile (PUC) and Universidad Austral de Chile (UACh), two universities of Ecuador: Universidad de Cuenca (UCuenca) and Escuela Superior Politécnica del Litoral (ESPOL) and three European universities: Universidad Carlos III de Madrid (UC3M), University of Edinburgh (UEdin) and KU Leuven (KUL) will be the members of the project consortium.

2 PROPOSAL DESCRIPTION

The main objective of the project is to improve the quality, efficiency and relevance of Higher Education in Latin America. Only by analyzing the different processes involved in higher education, the academic decision makers could understand and optimize these processes for the benefit of students and society. This project aims to build local capacity in in LatAm HEIs to create, adapt, implement and adopt Learning Analytics tools to improve academic decision making processes. These tools will facilitate the process and analysis of large amount of data produced by the different educational process that occur inside LatAm HEIs (registration, academic performance, online systems usage, etc.). Decision makers and stakeholders will use the output of these tools to inform and support their decisions. These evidence-based decision making process is bound to improve the performance and quality of the education inside the HEI.

To be able to develop the local capacity to create, adapt and use Learning Analytics tools in Latin America, we have defined four important milestones in order to achieve the main goal. These milestones are:

(1) A framework that describe the methodological, technical, institutional, ethical and communal aspects of the deployment of Learning Analytics in the context of LatAm HEIs should be developed by the project. The project will follow the example of what the SHEILA project did in Europe to gather information and opinions from key stakeholders in order to propose such contextualized framework.
(2) To test that the local technical capabilities are in place, the project will adapt two existing tools created originally in the context of Europe to the LatAm context. One will be directed to academic administrators and the other, to professors and counselors. Both of these tools will be piloted to test their efficacy to improve academic decision-making processes.
(3) The final result of the project will be the compilation of the LALA Handbook, a guide containing the LALA methodological framework, the Infrastructure recommendation and the adoption experiences and best-practices gained during the pilot. This handbook will be the guiding resources to any other LatAm HEI interested in adopting Learning Analytics to modernize their operations.
(4) During the project, a community will be form to continue the efforts of the project and to disseminate and exploit its outcomes. The LALA Community will serve as a communication channel to share experiences and tools after the project finishes.

3 METHODOLOGY

To ensure that the project will acomplish the milestones and the main goal, it was organized in a set of five stages: preparation, development, quality plan, dissemination and explotation, and management.

3.1 Phase 1: Preparation

The project will start with a preparation phase to ensure the smooth flow of the progress of the whole project. This phase has two main parts:

a) Set up of the project partners: that consist in organize and setup tools to be able to work together in the project. This selecting the working teams in each institution, forming the steering committee and having a first face-to-face meeting (Kick-off meeting).

b) Set up of the LALA Community: an inventory of HEIs that have experimented or are interested in the adoption of Learning Analytics tools will be created. To achieve this, an invitation will be sent to the identified instituions to join the LALA Community forming meeting that will be organized by the project.

3.2 Phase 2: Development

This phase will focus on the development of a framework to facilitate the design, implementation and use of Learning Analytics in the context of LatAm HEIs. This phase has three main parts:

a) Set up the LALA framework: The LALA Framework will provide guidelines to facilitate the design, implementation and use of Learning Analytics tools to improve decision making in academic processes in the context of the LatAm HEIs. To create this framework there will be a set of meetings and remote work that will bring together members of the project, academic decision makers, professors and students to interchange ideas, opportunities and barriers for the implementation of Learning Analytics tools in LatAm institutions, based on a participatory design strategy. The methodology set by the SHEILA project in Europe will be contextualized to gather the opinions and information from all relevant stakeholders: First, there will be a systematic review of existing policies and adoption studies in LatAm higher education. Second, a group concept mapping involving experts in learning analytics. Third, interviews with leaders of higher education institutions in Latin America. Finally, two sets of surveys, one measuring institutional readiness and another directed to students and teaching staff.

b) Adaptation of LALA tools: Two Learning Analytics tools will be adapted and contextualized its guidelines. First, a dashboard application will be adapted from the results of ABLE project. The ABLE dashboard is currently used at KU Leuven to support student counseling sessions: both the student and the study advisor see an overview of results obtained throughout the program and can compare these results to results of other students. In LALA, this dashboard will be taken as a starting point. The dashboard will be re-designed to address the requirements and needs of Latin-American universities. The design and development process used in ABLE project will be replicated: The first step will be elaborate interviews with faculty members and study advisors to understand the needs and issues that need to be addressed by the dashboard. In a second step, a first prototype will be elaborated that will be evaluated with study advisors. In a third step, the prototype will be improved based on feedback of from study advisors. Think aloud sessions will be conducted with study advisors to identify potential usability issues and refine the design.

The second tool consist of an adaptation tools created in UEdin and UC3M to serve as Drop-out Early Warning Systems. In a first step, the project will elaborate interviews will be conducted with academic program coordinators. Input will be collected on the different data sources that can be used, as well as metrics that should be visualized in the dashboard. Design goals will be defined in this first step. In a second step, a first prototype will be developed that addresses these design goals. Data sources will be collected and automatic analysis techniques will be deployed that provide useful insight. In a third step, the prototype will be tested with academic program coordinators. A think aloud study will be conducted to assess the usability and utility of the dashboard. This study will be used to identify potential issues and to refine the dashboard.

c) Piloting: The two Learning Analytics tools will be tested during a Piloting stage. During this stage, these tools will be integrated in the academic process of the LatAm HEIs. The Counseling Dashboard (CD) and the Drop-out Early Warning System (DEWS) will be integrated in the counseling sessions with students. By the end of the project, all the LatAm partners use this system as a regular tool for advisors or counselors. During all this process, data will be collected to evaluate the usefulness and impact of the Learning Analytics tools. The feedback obtained from the piloting will be used to improve both the LALA Framework. This new document, together with the description of the Learning Analytic tools (CD and DEWS) and the experiences gained during the piloting, will be compiled into the main outcome of the project, the LALA Handbook. This Handbook will be a guide to any other LatAm HEI to facilitate the adoption of Learning Analytics tools to improve their decision-making processes.

3.3 Phase 3: Quality plan

The purpose of this phase is to assure that the project has its expected impact and the activities fulfill their outcome as planned. The first part of this phase consists on the regular evaluation by independent party of the outcomes of the project. The external evaluation will be conducted by external experts once a year and its scope will be the results and impact of the project. These external experts will provide useful feedback to improve the project during its execution. The specific objective of the external evaluation will depend on the year when it is executed. During the first year, the external evaluation will be focused on the main outcomes of the project (LALA Framework and tools) to identify design flaws or opportunities for improvement. During the second year, the external experts will evaluate the piloting of the project to provide information about the impact on professors and students. At the end of the project, the external experts will evaluate the final impact of the project on LatAm HEIs. Second, as part of the Quality Assurance process, this phase also includes the design, update and enforcement of the Quality Plan. The purpose of these activities is to assure that all the activities and their outcomes reach an agreed level of quality and contribute to the success of the project.

3.4 Phase 4: Dissemination and Explotation

Different activities, directed at different stakeholders will be conducted during the project. First, the conceptual, technical and methodological part of the project will be discussed at Educational Conferences to increase the visibility of the project and obtain valuable early feedback from educational experts in the region. Another approach to disseminate the result of the project will be a continuous Social Media campaign (including social network sites such as Twitter, Facebook, YouTube, etc.) to reach the general public and raise awareness in the society. To help with the dissemination and exploitation of the project results, each partner will organize national training days where professors will be trained in the LALA Framework for the creation of their own Learning Analytic tools.

Finally, to help with the continuity of the LALA Community that will provide sustainability to the project, two LALA Community events will be organized during the project. The objectives of these events will be to showcase experiences and best-practices from academic and technical members of the LALA Community and attract new members from the participating public.

3.5 Phase 5: Management

The management of the project, will assure that the work in other packages runs smoothly and that any issue or conflict that arises during the execution of the project is solved. Management meetings will be held regularly. These meetings will be held face-to-face during other planned project meetings or virtually through the communication tools of the project. During the management meetings, the Steering Committee will review the status of the project and set the goals to be met in the next period. Also, the reached outputs will be analyzed and the Quality Plan will be update. This work package is also responsible for generating for the periodical project reports.

4 IMPACT AND EXPECTED RESULTS

The main impact expected from the project is the creation of local capacity to create, adapt, implement and adopt Learning Analytics tools to improve the academic decision-making process on LatAm HEIs. The outputs of the project will be used by four main beneficiaries: academic decision-makers, faculty, students and academic ICT providers. Decision-makers and faculty will use the LALA Framework to plan and design tools to help understand and optimize diverse educational processes. Academic ICT providers will also use the technical part of the LALA Framework to build interoperable tools that are adjusted to the needs of Decision-makers and faculty. Students will be the final beneficiaries of the improved decision-making processes, such as counseling or early feedback systems.

During the project, all those four stakeholders will be approached in the participating LatAm HEIs. Faculty and academic decision-makers will be reached through institutional and National training days and they will be involved in the adaptation of the LALA tools. Students will be reached to the application of the pilots in counseling sessions. Also, during the project, workshop events will be organized at the national and regional level to disseminate the results of the project to professors and academic authorities from other universities. Educational experts in Latin America will be reached through educational conferences presentations. Also we expect at least 8 large institutions in Latin America that regularly use Learning Analytics tools to take informed decisions, at least 2 Learning Analytics Tools has been adapted and developed in the context of the project by the LatAm HEis, at least 300 decision-makers/faculty involved in the pilots, at least 5000 students involved in the pilots, at least 120 additional professors decision-makers trained and at least 6 presentations at Educational Conferences.

The main beneficial effects of the project will be: 1) At local level, institutions will improve their decision-making process through tools that facilitate informed decisions, 2) At Regional level, the LALA Community will have a networking effect to share best practices and experiences in the use of Learning Analytics in LatAm HEIs, 3) At an institutional level, a methodological framework would be available to adopt tools for the improvement of learning processes in higher education on a large scale, 4) At European level, the results of the project will help European HEIs to fine-tuned their own use of Learning Analytics and will provide new partners to further explore the field.
As expected results of the project, we have definided the following:

(1) Create a LALA Community with interested LatAm HEIs
(2) Propose a LALA Framework to facilitate the adoption of Learning Analytics tools in LatAm HEIs. This framework will be created using the SHEILA methodology.
(3) Adapt 2 Learning Analytics tools (Counseling Tools from ABLE project and Drop-out Early Warning system from the UEdin and UC3M) to the LatAm context.
(4) Pilot the 2 tools in the 4 LatAm partners
(5) Create the LALA Handbook out of the experiences of the project.
(6) Disseminate the project through Conference presentations, National training days, Regional workshops (ECTEL and LACLO) and LALA Community.

5 CONCLUSIONS

Given the important technological advances, HEIs are able to access an important amount of stored data that represent the way in which the teaching and learning process has taken place in the educational programs they offer. The analysis of these data helps to make better decisions to adopt strategies that allow the effectiveness not only of the institution, but also the effectiveness in the training process since the use of LA can enable the structuring of learning through a personal learning environment.

Along these lines, this paper summarises the LALA project for building local capacity in LatAm HEIs to design, implement and use Learning Analytics tools to support their decision making processes. The project is conformed by different HEIs from LatAm and Europe, which combine knowledge and experience: Pontificia Universidad Católica de Chile (PUC), Universidad Austral de Chile (UACh), two universities of Ecuador: Universidad de Cuenca (UCuenca) and Escuela Superior Politécnica del Litoral (ESPOL) and three European universities: Universidad Carlos III de Madrid (UC3M), University of Edinburgh (UEdin) and KU Leuven (KUL) will be the members of the project consortium.

The aims of the project rely on the possibility to develop the local capacity to create, adapt and use Learning Analytics tools in Latin America through the development of a framework that describe the methodological, technical, institutional, ethical and communal aspects of the deployment of Learning Analytics. Also, we will adapt two existing tools for the context of LatAm HEIs. As a result, a guide containing the LALA Methodology, the Infrastructure recommendation and the adoption experiences and best-practices gained during the pilot will be published in the LALA Handbook. During the project, the

LALA community will be built to continue the efforts of the project and to disseminate and exploit its outcomes.

ACKNOWLEDGMENTS

This research project is co-funded by the European Commission Erasmus+ program, in the context of the project “Building Capacity to Use Learning Analytics to Improve Higher Education in Latin America”. The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. Please visit our website https://www.lalaproject.org/.

Make your courses count!: Using a digital analytics measurement model to measure the success of online courses

Chris Jennings
Google LLC
cjennings@google.com

ABSTRACT: In this half-day workshop, the team lead from Google's *Analytics Academy* will demonstrate how to create a digital analytics measurement model to evaluate online courses and learning programs. Based on the success of the Academy's own open learning environment, this workshop will walk participants through the structure of a digital analytics measurement model, how to decide on what metrics and dimensions should be prioritized for goals, how to create an implementation strategy to collect the right data, and work through the process of analyzing that data to understand whether the course and program has achieved its stated objectives. Once participants are familiar with how to create a digital analytics measurement model, they will have an opportunity to work collaboratively to create their own measurement model for their course/program and have it vetted by the group. The workshop will then discuss how to use the data analysis findings from the measurement model to communicate the successes and opportunities of courses and learning programs to stakeholders.

Keywords: course measurement, digital analytics, MOOCs, online courses, learning metrics

Workshop background

When speaking at conferences, I'm always surprised by the number of educators who have been inundated with messaging around the importance of big data and learning analytics, but have very little knowledge about what data to collect, how to analyze and make meaning of the data, and how to communicate the results to stakeholders. Google's *Analytics Academy* approaches learning analytics from the perspective of creating a digital analytics measurement plan for businesses. They use the structure and language of business measurement to articulate a useful framework for collecting learning data and understand whether a learning experience was successful. This is a different take on learning analytics with hands-on, practical application for all educators interested in making more data-driven strategic decisions about their curriculum. This topic has been presented at previous conferences to great interest and received an extremely positive reception. While previously presented in the context of a 45-minute presentation followed by Q&A, this submission proposes to turn this presentation into a hands-on workshop where participants create their own measurement plans to use as a template in their own programs and courses.

Workshop proposal

This event is proposed as a half-day workshop for participants. It is intended to be an open workshop, but should be limited to (ideally) 30 participants in order to maximize the quality of dialog and collaboration, and give the participants time to work together, present, and get feedback accordingly. The participants should expect some presentation from the speaker that lays down the fundamentals of digital analytics measurement, how a measurement plan works from a business perspective, and a look at how the Analytics Academy created a measurement model and regularly evaluates its own program and courses. Then participants will break into groups to develop their own measurement models, and present those models for feedback and discussion. For this workshop, a laptop and projector is all that's needed for the presenter. Participants should plan to bring their laptops to engage with the digital document templates that will be provided.

Below are details of the main sections of the proposed workshop:

1 The importance of building a learning measurement model

Learning measurement models are incredibly important when first designing an online course. These help associate finite metrics back to program objectives and course learning goals, and should be articulated prior to writing and building out a course. A learning measurement model will force educators to define their overall program objectives and course strategies, and connect those objectives and strategies back to measurable data. It will provide a guide for implementation and data collection, including any custom data required. It will also help define segments and custom dimensions to analyze data, as well as set goals for a learning program and courses. Additionally, it provides the metrics in a modality to tell the story of what worked in a course and what needs improvement.

2 Understanding the structure of a measurement plan

The idea of Google's *Analytics Academy* learning measurement model was adapted from digital analytics measurement plans used in online business. This is a method that allows businesses getting started with online data collection and analysis to define their business goals and objectives, and collect the right data to ensure they adequately measure their business goals and marketing. We'll walk through an example measurement plan for a small, non-profit business, exploring the various parts of a measurement plan such as business objectives, online strategies, deployment tactics, key performance indicators (or KPIs), segments for slicing data for analysis, and goals.

3 Exploring the Academy learning measurement model

Next, we'll look at the Analytics Academy learning measurement model and walk through the objectives, strategies, tactics, KPIs, etc. for the program. We'll look at how the Academy tied its measurement and data collection back to specific program and business goals, and walk through all of the various KPIs that are used to drive analytics implementation and measure course success. We'll break this down into knowledge metrics, engagement metrics, sentiment, and product engagement, as well as a number of dimensions and Custom Dimensions used to better understand audience characteristics and motivation. We'll also dispense strategies and advice for setting realistic goals, ensuring data quality, ongoing evaluation of the data, and best practices.

Performing analysis and evaluating learning data against goals

Google's Analytics Academy will then walk participants through their method of data analysis, how data was collected and formatted for presentation, how they interpreted the data that came back, and how that influenced the program/course strategically. We'll look at the difficulty of particular kinds of data collection, how to collect data from disparate systems, and how to create dashboards that help analyze data on an ongoing basis.

4 How to build a learning measurement model

Based on the concepts explained for designing a learning measurement model, the workshop participants will break into groups and use worksheets to help design their own learning measurement model that will articulate data collection and metrics for their own program/course(s). This will be an interactive session where participants can ask questions, seek collaborative help from others, and ask specific questions of the presenter. Each group will then have an opportunity to explain their measurement model, discuss some of the decisions that were made it creating the model, and what different results might tell them about their program/course effectiveness.

5 Telling the story of program/course success and opportunities to improve

The workshop will close with advice about telling stories from your learning data and communicating data-driven decisions. This will instruct participants in how to take the findings from their data analysis and use that to describe course successes and opportunities. Creating dashboards and decks that show how well your courses worked and are honest and upfront about iterative improvement, can help influence stakeholders and provide ongoing support for programs, while helping communicate a confident strategic direction based on data that ties back to learning and program goals.

Workshop objectives and intended outcomes

This workshop aims to give participants an introduction to the specifics of digital analytics measurement including metrics, dimensions, Custom Dimensions, data collection strategies, and ideas around how to measure specific aspects of online education (particularly at scale, using the Analytics Academy MOOC example). It will walk users through the parts of a measurement plan, how to articulate the various ways a program/course will be measured, and some sample data analysis. It will then empower each participant to create their own measurement model for real-world use and provide examples and advice on how to demonstrate those results to stakeholders. Each participant should leave the session with a fully-articulated learning measurement model that can be implemented and used to make data-driven decisions about their learning experiences.

References

Dr. Chris Dede Professor, Harvard University chris_dede@gse.harvard.edu

Dr. Darcy Hardy Associate Vice President, Enterprise Consulting at Blackboard darcywhardy@gmail.com

Les Howles Director, Distance Education Professional Development, Division of Continuing Studies at the University of Wisconsin – Madison les.howles@wisc.edu

Dr. Scott M. Martin Founding Director, Virginia Serious Game Institute (VSGI) Associate Professor, Computer Game Design at George Mason University smartin4@gmu.edu

Dr. Rob Robinson Senior Director, Strategic Services at Civitas Learning at Civitas Learning rob.robinson@civitaslearning.com

Justin Cutroni Director, Google Analytics Education at Google jcutroni@google.com

Participatory Design of Learning Analytics (PD-LAK)

Carlos Prieto Alvarez[1], Theresa Anderson[2], Roberto Martinez-Maldonado[3], Kirsty Kitto[4]
University of Technology Sydney
CarlosGerardo.PrietoAlvarez@student.uts.edu.au[1], TheresaAnderson@uts.edu.au[2],
Roberto.Martinez-Maldonado@uts.edu.au[3], Kirsty.Kitto@uts.edu.au[4]

Mollie Dollinger
The University of Melbourne
Mollie.Dollinger@unimelb.edu.au,

Jen McPherson
Macquarie University
Jen.Mcpherson@mq.edu.au

1 BACKGROUND AND MOTIVATION

There is an increasing interest in the use of learning analytics by individual educators and at different institutional levels (Daniel, 2017). However, researchers and designers may be challenged to understand different learning settings, student motivation, and stakeholder expectations (Drachsler & Greller, 2012). From a learner-centred design perspective, these problems can, to a certain extent, be explained in terms of the lack of student and educator involvement in the design of the tools that are intended to support their learning and/or teaching (DiSalvo et al., 2017).

Participatory Design can complement traditional design to improve current learning analytics systems. Incorporating qualitative approaches from Participatory Design with quantitative methods used in educational data mining, creates potential to design better learning analytics innovations. Some effort has been made to include learners and other stakeholders in learning analytics design, for example in setting privacy and ethical policies (Slade & Prinsloo, 2015), understanding students' perspectives on data and learning analytics systems (McPherson et al., 2016) and stakeholders' expectations towards their learning analytics tools (Drachsler & Greller, 2012). One limitation of these attempts is that it is not clear how their approaches can be replicated so other researchers and designers can apply techniques to Participatory Design of learning analytics tools.

Bringing researchers and practitioners to this area requires a shared understanding of how Participatory Design works in this context, which can be developed by bringing interested practitioners together with current practitioners in this field. Similar workshops in other education and technology areas have been conducted in recent years indicating an interest for using Participatory Design within learning analytics. Some examples include Andrews et al. (2014); Anthony et al. (2012); and a hand on session such as Anderson and Knight (2016). Running a workshop like this for the LAK community may lead to further collaboration between researchers and better practices for people interested in using Participatory Design tools and techniques.

2 GOALS AND THEMES

This workshop focuses on Participatory Design for learning analytics. The main goals for this workshop are i) introducing participants to core principles of co-design that can be applied to LA contexts; ii) providing participants with opportunity to sample some tools, techniques, methods; iii) help participants get more comfortable with the challenges associated with such open-ended, participatory approaches to designing LA objects/projects; iv) foster the development of a community of Participatory Design for LA practice. Participants attending this workshop will be able to take home ideas and material provided as a first step into implementing Participatory Design in their projects.

This workshop aims to generate understanding over participatory design techniques following three themes from benefits, opportunities and current practices to actual implementation:

1. Participatory Design in education: Case studies and examples of application of Participatory Design practices in educational settings where learning analytics are intended to be adopted by teachers and learners.
2. Tools and techniques for Participatory Design: Examples of Participatory Design tools and techniques suitable for adoption by the LA community.
3. Participatory Design challenges and future implementation in LA: Understand the challenges and design principles as a first step in the Participatory Design process. This also includes lessons learned from current practitioners in the field.

3 OUTCOMES

The intended outcomes of this workshop are to increase understanding of Participatory Design in learning analytics design, facilitate networking with current Participatory Design practitioners, help participants with their current projects through sharing guidelines, tools and techniques, kickstart a community of Participatory Design for LA practice. Participants will learn about current approaches to co-design and use this information to build, enhance, and update their own projects, including dashboards, LMS plugins, recommender systems and dashboards.

Papers presented in this workshop reflect the current interest from designers, researchers and practitioners into using a participatory approach to current projects. Activities planned for this session focuses on the sharing of participants' experiences in designing with data. Organizers encourage participants to bring their personal cases and data sets to work with during scheduled activities. Activities planned aim to develop further our goals explained in the last section marked as i, ii, iii and iv. Outcomes, stories and material used for the session will be uploaded to workshop's website: *pdlak.utscic.edu.au.*

REFERENCES

Anderson, T., & Knight, S. (2016). Participatory Design of a Learning Space and Learning Analytics. *Teaching events.* Retrieved from https://teaching.unsw.edu.au/participatory-design

Andrews, B., Bardzell, S., Clement, A., D'Andrea, V., Hakken, D., Poderi, G., . . . Teli, M. (2014). *Teaching participatory design*. Paper presented at the Proceedings of the 13th Participatory Design Conference: Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium papers, and Keynote abstracts - Volume 2, Windhoek, Namibia.

Anthony, L., Prasad, S., Hurst, A., & Kuber, R. (2012). *A participatory design workshop on accessible apps and games with students with learning differences*. Paper presented at the Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility, Boulder, Colorado, USA.

Daniel, B. K. (2017). Overview of Big Data and Analytics in Higher Education. In B. Kei Daniel (Ed.), *Big Data and Learning Analytics in Higher Education: Current Theory and Practice* (pp. 1-4). Cham: Springer International Publishing.

DiSalvo, B., Yip, J., Bonsignore, E., & DiSalvo, C. (2017). *Participatory Design for Learning: Perspectives from Practice and Research*: Taylor & Francis.

Drachsler, H., & Greller, W. (2012). *The pulse of learning analytics understandings and expectations from the stakeholders*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.

McPherson, J., Tong, H. L., Fatt, S. J., & Liu, D. (2016). *Student perspectives on data provision and use: starting to unpack disciplinary differences.* Paper presented at the Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, NY, USA.

Slade, S., & Prinsloo, P. (2015). Student perspectives on the use of their data: between intrusion, surveillance and care. *Student perspectives on the use of their data: between intrusion, surveillance and care*, 291-300.

Collaborative Personas for Crafting Learners Stories for Learning Analytics Design

Carlos Prieto Alvarez, Vanessa Echeverria Barzola, Theresa Anderson, Roberto Martinez-Maldonado
University of Technology Sydney
CarlosGerardo.PrietoAlvarez@student.uts.edu.au,
Vanessa.i.echeverriabarzola@student.uts.edu.au, TheresaAnderson@uts.edu.au,
Roberto.Martinez-Maldonado@uts.edu.au

ABSTRACT: Learning analytics innovations are attracting the attention of researchers and designers for providing personalized learning experiences, tracking improvement and better understanding the social aspects of learning. However, current design practices often neglect learners' involvement causing a misalignment with learners' intentions and representations of learners. In this paper, we explore the role of current user modeling tools such as Persona Profile in a collaborative setting to enhance representations of learners This illustrative case study describes the main finding when designing representation with current students in the Bachelor of Nursing program.

Keywords: Learning Analytics, User Modeling, Participatory Design

1 INTRODUCTION AND RELATED WORK

Design for Learning analytics (LA) is a new field where often interdisciplinary teams work together including designers, tutors and researchers. When it comes to define who is the main beneficiary, designers often find that representing learners is an elaborated task that requires lots of effort. In particular for learning analytics, researchers are starting to bring students along the design process including this particular steep that consists of stablishing roles and using this information further down the co-design process (Jisc, 2016) (McGregor, 2016a).

User representations became popular as a design object when researchers from marketing areas tried to enhance classic segmentation(Cooper, Reimann, Cronin, & Noessel, 2014), in terms of learning, researchers and designers may benefit from bringing this into LA adapting the original application to learning scenarios. In design for learning, practitioners use representations to describe what a learner is. With a traditional implementation, we still find misalignment on what constitutes a learner from the designers' perspective and how learners see themselves in the design for learning ecosystem (McGregor, 2016a)

The structure of this papers starts with a brief description of how a widely used tool from design practices like Personas can be used in a participatory setting to enhance representations of learners. The following sections describe an illustrative study involving students from the Bachelor of Nursing program, the process followed during the participatory design sessions and a preliminary analysis on resulting data used for crafting User Stories.

2 PERSONA PROFILE AS LEARNERS REPRESENTATIONS

In the current field of design, we find some practical tools that help designers to create user representations from different sources of data. Designers used general characteristics to generate a first profile including age, gender, occupation and familiarity with technology (Junior & Filgueiras, 2005). Establishing what characteristics are most descriptive can be problematic and in learning settings, students may be subject for profiling and false assumptions (Goodwin, 2005).

Personas have proven to be useful when it comes to practical representations. There are few examples of how Persona representations are being used as practical objects aligning users' expectations with their goals and needs (Junior & Filgueiras, 2005). Still, there are some challenges when LA designers put together learners' characteristics from available information without falling into a biased confirmation for our profile (Marsden & Haag, 2016).

Collaborative design methods for crafting representations are starting to be an interesting approach to address the current challenges. Regardless of the field, collaboration is helping designers shape better Personas as pointed by (Nielsen & Hansen, 2014) and (LeRouge, Ma, Sneha, & Tolle, 2013). However, in education related areas such as Learning analytics, Learning design and Educational Data Mining there is still some work to do when it comes to representations in collaborative settings (Bodker, Christiansen, & Nyvang, 2012; McGregor, 2016b). Involving students in design sessions can bring an additional layer of data that can be enhanced with tutors and designers perspectives (Sciences, 2016) (IIDC, 2015). When designing learning innovations, these representations can be useful for generating a unified vision of who the educational technology will provide support to and how learners expect the product to align with their personal interests and goals (Gladys Castillo, Joao Gama, & Breda, 2006) (Pruitt & Adlin, 2005).

3 ILLUSTRATIVE STUDY

3.1 Learning analytics design context

During the Bachelor of Nursing program, students require to attend a series of practice-oriented courses involving simulations while using representative tools such as manikins, professional equipment and support machinery. The learning objective is to provide accurate scenarios where different skills are developed including communication (with patients and peers), problem solving and leadership. When learning a new procedure (examples Cardiopulmonary resuscitation (CPR), life support, stroke) students require to read/watch the relevant material provided by tutors, demonstrate in group work how to conduct this new practice and then reflect on what can be improve.

The challenge for us is to design an automated feedback tool to provide learners with useful insights on their practice. Using learning analytics tools, it is possible to track different sorts of activity around the classroom. In this initial part of the study, it is important to understanding how learners can help to create better representations in this scenario and what other problems may find benefit from using a data intensive approach such as learning analytics.

Learners' involvement in crafting their own representations in the whole scheme of design requires a set of tools and techniques tested to facilitate collaboration between learners, designers and tutors. In this case Personas designed in collaboration provides a collective view on what represents a learner and how they see themselves as users in this context.

3.2 Persona profile template

An initial template was crafted by the lead designer based on what tutors and researchers stablished as main interests. This template includes fields starting with what values are being endorse as nurses in training, learners' goals, and open topics that may not be expressed in other ways as students. Other practical fields include the media they prefer and what other activities they do after class.

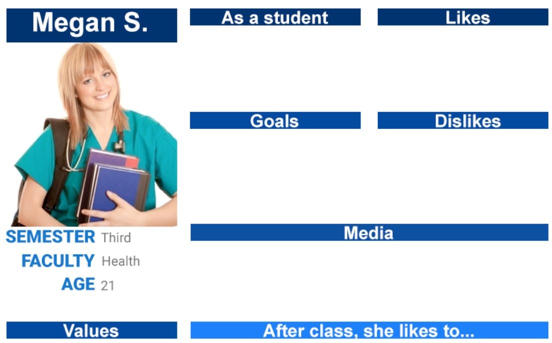


Figure 1: Template provided as an initial representation object.

A series of co-design sessions were conducted with students. The sessions were distributed with 15 (N=15) students across 5 group sessions (GS=5). The research team conducted a focus group approach with a guided scheme of activities. Learners were asked to use the template as an initial description to fill while conversations were recorded for further analysis.

3.3 Preliminary analysis

After conducting the sessions, we gathered and compared the different profiles built by participants. Field and notes were added to our initial template based on learners' feedback and observations. The new fields added were on specific goals for the simulations, the different uses of social media and the reason behind wanting to become a nurse.

In table 1 we describe observations gathered from the conversations during the activity and new fields requested based on learners' feedback.

Group	Observations	New fields
1	• Personal and global values are hard to express in one single field. • Media sources used by learners differs based on technology expertise.	• Personal goals and academic goals. • Social media and LMS preferences.
2	• After class activities can be used to express leisure and additional hobbies.	• A field for open comments on personal traits.
3	• Values generates discussion since this term is not used by tutors or any resource provided by the faculty.	• Concerns.
4	• Academic goals are different from personal goals.	• Academic goals.
5	• A different template for seniors and new students.	• Current challenges.

Table 1- Observations and new fields requested by leaners per session.

To continue with our study, some recommendations are being followed based on these preliminary results. The first one is to complement our Persona profiles with comments provided by learners and show this to the design team. Some components on the template are mere suggestion than designer may not followed but now they become aware in case additional information back up the suggestion. After a further analysis on how the session was conducted and current observations from participants, we describe two main recommendations that may help researchers to improve representations of learners in collaboration.

Encourage discussion and decision: When writing something down in the template encourage learners to decide at least three terms and ask them to explain why that information is relevant.

Template customization: Diversified groups may require a different template based on the group composition. Changes on the template between session can make the analysis process more challenging but allows to gather more data that can reviewed further down the design process.

4 CONCLUSION

Building collaborative representations by using a Persona template helps researchers and designers to open the design process to learners. The resulting objects can be used in the future to generate usage scenarios where learning analytics innovations can be deployed. Also, these objects become a resource for designers and other stakeholders to comprehend user intentions without going into technicalities. For the following sessions, some other techniques from PD and Co-design areas will be tested to gather additional information and support collaboration through the whole design process of LA tools.

REFERENCES

Bodker, S., Christiansen, E., & Nyvang, T. (2012). *Personas, people and participation: challenges from the trenches of local government*. Paper presented at the Proceedings of the 12th Participatory Design Conference: Research Papers - Volume 1, Roskilde, Denmark.

Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About Face: The Essentials of Interaction Design*: Wiley.

Gladys Castillo, Joao Gama, & Breda, A. (2006). Adaptive Predictive Model for Student Modeling. In Idea Group Inc (IGI) (Ed.), *Advances in Web-based Education: Personalized Learning Environments* (pp. 371): Idea Group Inc

Goodwin, K. (2005). Perfecting Your Personas. Retrieved from https://www.cooper.com/journal/2008/05/perfecting_your_personas

IIDC. (2015). Student Personal Profile *IIDC Design Resources.* Retrieved from https://www.iidc.indiana.edu/styles/iidc/defiles/INSTRC/Student_Personal_Profile_Template.pdf

Jisc. (2016). Effective learning analytics. *Effective Learning Analytics.* Retrieved from https://www.jisc.ac.uk/rd/projects/effective-learning-analytics

Junior, P. T. A., & Filgueiras, L. V. L. (2005). *User modeling with personas*. Paper presented at the Proceedings of the 2005 Latin American conference on Human-computer interaction, Cuernavaca, Mexico.

LeRouge, C., Ma, J., Sneha, S., & Tolle, K. (2013). User profiles and personas in the design and development of consumer health technologies. *International Journal of Medical Informatics, 82*(11), e251-e268. doi:https://doi.org/10.1016/j.ijmedinf.2011.03.006

Marsden, N., & Haag, M. (2016). *Stereotypes and Politics: Reflections on Personas*. Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, California, USA.

McGregor, A. (2016a). Co-design-consultation 2016-17. *R&D.* Retrieved from https://www.jisc.ac.uk/rd/how-we-innovate/co-design-consultation-2016-17

McGregor, A. (2016b). Codesign challenge: Here be data monsters. *Jisc student experience blog.* Retrieved from https://elearning.jiscinvolve.org/wp/2016/11/15/codesign-challenge-here-be-data-monsters/

Nielsen, L., & Hansen, K. S. (2014). *Personas is applicable: a study on the use of personas in Denmark*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada.

Pruitt, J., & Adlin, T. (2005). *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*: Morgan Kaufmann Publishers Inc.

Sciences, A. (2016). Fictional Student Personas *Communications and Marketing.* Retrieved from http://agsci.psu.edu/communications/audiences/general/undergraduate-student/personas

Context-Appropriate Scaffolding Assemblages: A generative learning analytics platform for end-user development and participatory design

David Jones
University of Southern Queensland
david.jones@usq.edu.au

Celeste Lawson
CQUniversity
c.lawson@cqu.edu.au

Colin Beer
CQUniversity
c.beer@cqu.edu.au

Hazel Jones
University of Southern Queensland
hazel.jones@usq.edu.au

ABSTRACT: There remains a significant tension in the development and use of learning analytics between course/unit or learning design specific models and generic, one-size fits all models. As learning analytics increases its focus on scalability there is a danger of erring toward the generic and limiting the ability to align learning analytics with the specific needs and expectations of users. This paper describes the origins, rationale, and use cases of a work in progress design-based research project attempting to develop a generative learning analytics platform. Such a platform encourages a broad audience to develop unfiltered and unanticipated changes to learning analytics. It is hoped that such a generative platform will enable the development and greater adoption of embedded and contextually specific learning analytics and subsequently improve learning and teaching. The paper questions which tools, social structures, and techniques from participatory design might inform the design and use of the platform, and asks whether or not participatory design might be more effective when partnered with generative technology?

Keywords: Contextually Appropriate Scaffolding Assemblages (CASA); generative platform; participatory design; DIY learning analytics

1 INTRODUCTION

One size does not fit all in learning analytics. There is no technological solution that will work for every teacher, every time (Mishra & Koehler, 2006). Context specific models improve teaching and learning, yield better results and improve the effectiveness of human action (Baker, 2016; Gašević, Dawson, Rogers, & Gasevic, 2016). Despite this, higher education institutions tend to adopt generalised approaches to learning analytics. Whilst this may be cost effective and efficient for the organisation (Gašević et al., 2016), the result is a generic approach that provides an inability to cater for the full diversity of learning and learners and shows "less variety than a low-end fast-food restaurant" (Dede, 2008).

Institutional implementation of learning analytics in terms of both practice and research remain limited to conceptual understandings and are empirically narrow or limited (Colvin, Dawson, Wade, & Gašević, 2017). In practice, learning analytics has suffered from a lack of human-centeredness (Liu, Bartimote-Aufflick, Pardo, & Bridgeman, 2017). Even when learning analytics tools are designed with the user in mind (e.g. Corrin et al., 2015), the resulting tools tend to be what Zittrain (2008) defines as non-generative or sterile. In particular, the adoption of such tools tends to require institutional support and subsequently leans toward the generic, rather than the specific. This perhaps provides at least part of the answer of why learning analytics dashboards are seldom used to intervene during the teaching of a course (Schmitz, Limbeek, van Greller, Sloep, & Drachsler, 2017) and leading us to the research question: How can the development of learning analytics better support the needs of specific contexts, drive adoption, and ongoing design and development? More broadly, we are interested in if and how learning analytics can encourage the adoption of practices that position teaching as design and subsequently improve learning experiences and outcomes (Goodyear, 2015) by supporting a greater focus on the do-it-with (DIW - participatory design) and do-it-yourself (DIY) design (where teachers are seen as designers), implementation, and application of learning analytics. This focus challenges the currently more common Do-It-To (DIT) and Do-It-For (DIF) approaches (Beer, Tickner & Jones, 2014).

This project seeks to explore learning analytics using a design-based research approach informed by a broader information systems design theory for e-learning (Jones, 2011), experience with Do-It-With (DIW) (Beer et al., 2014) and teacher Do-It-Yourself (DIY) learning analytics (Jones, Jones, Beer, & Lawson, 2017), and technologies associated with reproducible research to design and test a generative learning analytics platform. Zittrain (2008) defines a generative system as having the “capacity to produce unanticipated change through unfiltered contributions from broad and varied audiences” (p. 70). How generative a system is depends on five principal factors: (1) *leverage*; (2) *adaptation*; (3) ease of *mastery*; (4) *accessibility*; and (5) *transferability* (Zittrain, 2008). A focus for this project is in exploring how and if a generative learning analytics platform can act as a boundary object for the diverse stakeholders involved with the design, implementation and use of institutional learning analytics (Suthers & Verbert, 2013). Such an object broadens the range of people who can engage in creative acts of making learning analytics as a way to make sense of current and future learning and teaching practices and the contexts within which it occurs. The platform - named CASA, an acronym standing for Contextually Appropriate Scaffolding Assemblages - will be designed to enable all stakeholders alone or together to participate in decisions around the design, development, adoption and sharing of learning analytics tools. These tools will be created by combining, customising, and packaging existing analytics - either through participatory design (DIW) or end-user development (DIY) - to provide context-sensitive scaffolds that can be embedded within specific online learning environments..

2 KNOW THY STUDENT – TEACHER DIY LEARNING ANALYTICS

Jones et al., (2017) uses a case of teacher DIY learning analytics to draw a set of questions and implications for the institutional implementation of learning analytics and the need for CASA. The spark for the teacher DIY learning analytics was the observation that it took more than 10 minutes, using two separate information systems including a number of poorly designed reports, to gather the information necessary to respond to an individual learner's query in a discussion forum. The teacher was able to design an embedded, ubiquitous and contextually specific learning analytics tool (Know Thy Student) that reduced the time taken to gather the necessary information to a single mouse click. The tool was used in four offerings of a third year teacher education unit across 2015 and 2016. Analysis of usage logs indicates that it was used 3,100 separate times to access information on 761 different students, representing 89.5% of the total enrolled students. This usage was spread across 666 days over the two years, representing 91% of the available days during this period. A significant usage level, especially given that most learning analytics dashboards are seldom used to intervene during the teaching of a course (Schmitz et al., 2017). Usage also went beyond responding to discussion forum questions. Since the tool was unintentionally available throughout the entire learning environment (embedded and ubiquitous) unplanned use of the tool developed contributing to improvements in the learner experience. This led to the implication that embedded, ubiquitous, contextual learning analytics encourages greater use and enables emergent practice (Jones et al., 2017). It provides *leverage* to make the difficult job of teaching a large enrolment, online course easier. However, the implementation of this tool required significant technical knowledge and hence is not easy to *master*, not *accessible*, nor easily *transferable*, Zittrain's (2008) remaining principles required for a generative platform. The questions now become: How to reduce this difficulty? *How to develop a generative learning analytics platform?*

3 CASA TECHNOLOGIES AND TECHNIQUES

To answer this question CASA will draw on a combination of common technologies associated with reproducible research including virtualisation, literate computing (e.g. Jupyter Notebooks), and version control systems (Sinha & Sudhish, 2016) combined with web augmentation (Díaz & Arellano, 2015) and scraping (Glez-Peña, Lourenço, López-Fernández, Reboiro-Jato, & Fdez-Riverola, 2014). Reproducible research technologies enable CASA to draw upon a large and growing collection of tools developed and used by the learning analytics and other research communities. Growth in the importance of reproducible research also means that there is a growing number of university teaching staff familiar with the technology. It also means that there is emerging research literature sharing insights and advice in supporting academics to develop the required skills (e.g. Wilson, 2016). Virtualisation allows CASA to be packed into a single image which allows individuals to easily download, install and execute within their own computing platforms. Web augmentation provides the ability to adapt existing web-based learning environments to embed learning analytics directly into the current common learning context. The combination of these technologies will be used to implement the CASA platform, enabling the broadest possible range of stakeholders to individually and collaboratively design and implement different CASA instances. Such instances can be mixed and matched to suit context-specific requirements and shared amongst a broader community. The following section provides a collection of CASA use case scenarios including explicit links to Zittrain's (2008) five principal factors of a generative platform.

4 CASA USE CASE SCENARIOS

A particular focus with the CASA platform is to enable individual teachers to adopt CASA instances while minimising the need to engage with institutional support services (*accessibility*). Consequently a common scenario would be where a teacher (Cara) observes another teacher (Daniel) using a CASA instance. It is obvious to Cara that this specific CASA instance makes a difficult job easier (*leverage*) and motivates her to trial it. Cara visits the CASA website and downloads and executes a virtual image (the CASA instance) on her computer, assuming she has local administrator rights. Cara configures CASA by visiting a URL to this new CASA instance and stepping through a configuration process that asks for some context specific information (e.g. the URL for Cara's course sites). Cara's CASA uses this to download basic clickstream and learner data from the LMS. Finally, Cara downloads the Tampermonkey browser extension and installs the CASA user script to her browser. Now when visiting any of her course websites Cara can access visualisations of basic clickstream data for each student.

To further customise her CASA instance Cara uploads additional data to provide more contextual and pedagogical detail (*adaptation*). The ability to do this is sign-posted and scaffolded from within the CASA tool (*mastery*). To expand the learner data Cara sources a CSV file from her institution's student records system. Once uploaded to CASA all the additional information about each student appears in her CASA and Cara can choose to further hide, reveal, or re-order this information (*adaptation*). To associate important course events (Corrin et al., 2015) with the clickstream data Cara uses a calendar application to create an iCalendar file with important dates (e.g assignment due dates, weekly lecture times). This is uploaded or connected to CASA and the events are subsequently integrated into the clickstream analytics. At this stage, Cara has used CASA to add embedded, ubiquitous and contextually specific learning analytics about individual students into her course site. At no stage has Cara gained access to new information. CASA has simply made it easier for Cara to access this information, increasing her efficiency (*leverage*). This positive experience encourages Cara to consider what more is possible.

Cara engages in a discussion with Helen, a local educational designer. The discussion explores the purpose for using learning analytics and how it relates to intended learning outcomes. This leads to questions about exactly how and when Cara is engaging in the learning environment. This leads them to engage in various forms of participatory design with Chuck (a software developer). Chuck demonstrates how the student clickstream notebook form Cara's existing instance can be copied and modified to visualise staff activity (*mastery*). Chuck also demonstrates how this new instance can be shared back to the CASA repository and how this process will eventually allow Daniel to choose to adopt this new instance (*transferability*). These discussions may also reveal insights into other factors such as limitations in Cara's conceptions and practices of learning and teaching, or institutional factors and limitations (e.g. limited quality or variety of available data).

5 CONCLUSIONS AND QUESTIONS

This paper has described the rationale, origins, theoretical principles, planned technical implementation and possible use cases for CASA. CASA is a generative learning analytics platform which acts as a boundary object. An object that engages diverse stakeholders more effectively in creative acts of making to help make sense of and respond to the diversity and complexity inherent in learning and teaching in contemporary higher education. By allowing both DIW (participatory design) and DIY (end-user development) approaches to the implementation of learning analytics we think CASA can enable the development of embedded, ubiquitous and contextually specific applications of learning analytics, better position teaching as design, and subsequently improve learning experiences and outcomes. As novices to the practice of participatory design we are looking for assistance in examining how insights from participatory design can inform the design and use of CASA. For us, there appear to be three areas of design activity where participatory design can help and a possibility where the addition of generative technology might help strengthen participatory design.

First, the design of the CASA platform itself could benefit from participatory design. A particular challenge to implementation within higher education institutions is that as a generative platform CASA embodies a different mindset. A generative mindset invites open participation and assumes open participation provides significant advantage, especially in terms of achieving contextually appropriate applications. It sees users as partners and co-designers. An institutional mindset tends to see users as the subject of design and due to concerns about privacy, security, and deficit models seek to significantly limit participation in design. Second, the DIW interaction between Cara, Helen and Chuck in the use case section is a potential example of using participatory design and the CASA platform to co-design and co-create contextually specific CASA instances. What methods, tools and techniques from participatory design could help these interactions? Is there benefit in embedding support for some of these within the CASA platform? Lastly, the CASA approach also seeks to enable individual teachers to engage in DIY development. According to Zittrain (2008) the easier we can make it for teachers to develop their own CASA instances (*mastery*) the more generative the platform will be. What insights from participatory design might help increase CASA's generative nature? Can CASA be seen as an example of a generative toolkit (Sanders & Strappers, 2014)? Or, does the DIY focus move into the post-design stage (Sanders & Strappers, 2014)? Does it move beyond participatory design? Is the combination of participatory design and generative technology something different and more effective than participatory design alone? If it is separate, then how can the insights generated by DIY making with CASA be fed back into the on-going participatory design of the CASA platform, other CASA instances, and sense-making about the broader institutional context?

REFERENCES

Baker, R. (2016). Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600–614. https://doi.org/10.1007/s40593-016-0105-0

Beer, C., Tickner, R., & Jones, D. (2014). Three paths for learning analytics and beyond : moving from rhetoric to reality. In B. Hegarty, J. McDonald, & S. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin 2014* (pp. 242–250).

Colvin, C., Dawson, S., Wade, A., & Gašević, D. (2017). Addressing the Challenges of Institutional Adoption. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The Handbook of Learning Analytics* (pp. 281–289). Alberta, Canada: Society for Learning Analytics Research.

Corrin, L., Kennedy, G., Barba, P. D., Williams, D., Lockyer, L., Dawson, S., & Copeland, S. (2015). Loop : A learning analytics tool to provide teachers with useful data visualisations. In T. Reiners, B. von Konsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), *Globally connected, digitally enabled. Proceedings ascilite 2015* (pp. 57–61).

Dede, C. (2008). Theoretical perspectives influencing the use of information technology in teaching and learning. In J. Voogt & G. Knezek (Eds.), *International Handbook of Information Technology in Primary and Secondary Education* (pp. 43–62). New York: Springer.

Díaz, O., & Arellano, C. (2015). The Augmented Web: Rationales, Opportunities, and Challenges on Browser-Side Transcoding. *ACM Trans. Web*, *9*(2), 8:1–8:30.

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicating learning success. *The Internet and Higher Education*, *28*, 68–84. https://doi.org/10.1016/j.iheduc.2015.10.002

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., & Fdez-Riverola, F. (2014). Web scraping technologies in an API world. *Briefings in Bioinformatics*, *15*(5), 788–797. https://doi.org/10.1093/bib/bbt026

Goodyear, P. (2015). Teaching As Design. *HERDSA Review of Higher Education*, *2*, 27–50.

Jones, D. (2011). An Information Systems Design Theory for E-learning (Doctoral thesis, Australian National University, Canberra, Australia). Retrieved from https://openresearch-repository.anu.edu.au/handle/1885/8370

Jones, D., Jones, H., Beer, C., & Lawson, C. (2017, December). *Implications and questions for institutional learning analytics implementation arising from teacher DIY learning analytics*. Paper presented at the ALASI 2017: Australian Learning Analytics Summer Institute, Brisbane, Australia. Retrieved from http://tiny.cc/ktsdiy

Liu, D. Y.-T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-Driven Personalization of Student Learning Support in Higher Education. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends* (pp. 143–169). Springer International Publishing.

Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054.

Sanders, E. B.-N., & Stappers, P. J. (2014). Probes, toolkits and prototypes: three approaches to making in codesigning. CoDesign, *10*(1), 5–14. https://doi.org/10.1080/15710882.2014.888183

Schmitz, M., Limbeek, E. van, Greller, W., Sloep, P., & Drachsler, H. (2017). Opportunities and Challenges in Using Learning Analytics in Learning Design. In Data Driven Approaches in Digital Education (pp. 209–223). Springer, Cham.

Sinha, R., & Sudhish, P. S. (2016). A principled approach to reproducible research: a comparative review towards scientific integrity in computational research. In *2016 IEEE International Symposium on Ethics in Engineering, Science and Technology (ETHICS)* (pp. 1–9).

Suthers, D., & Verbert, K. (2013). Learning analytics as a middle space. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13* (pp. 2–5).

Wilson, G. (2014). Software Carpentry: lessons learned. *F1000Research*, 3.

Zittrain, J. (2008). *The Future of the Internet--And How to Stop It*. Yale University Press.

Contextual Inquiry, Participatory Design, and Learning Analytics: An Example

Amna Liaqat[1], Benett Axtell[1], Cosmin Munteanu[1,2], Carrie Demmans Epp[3]
University of Toronto, TAGlab[1] | University of Toronto Mississauga, ICCIT | University of Alberta[3]
a.liaqat@mail.utoronto.ca, benett@taglab.ca, cosmin.munteanu@utoronto.ca, cdemmansepp@ualberta.ca

ABSTRACT: The methods used in learning analytics for early specification of design requirements are still generally grounded in prior research, theoretical frameworks, and the existing body of practice. These traditional methods provide a strong background for development, but adapting them to a wide range of user needs is challenging. Participatory design and contextual inquiry can address this challenge. These user-centred design methods help extend theoretical principles into real-world applications. As such, we propose field-based contextual inquiry and participatory design methods to elicit design requirements for learning analytics features and present an exemplar study as a starting point for future exploration and validation of these approaches.

Keywords: Contextual Inquiry, Participatory Design, Learning Analytics, Learning Dashboards

1 INTRODUCTION

Participatory design (PD) integrates users into the technology creation process through a variety of methods (e.g., interviews, observations, or design activities; Muller, 1993, 2003) to elicit requirements from the early stages of the design process. Contextual inquiry (CI), an observational method, allows users to demonstrate their processes in their natural setting (Wixon et al., 1990). Like in PD, a key component of CI is the partnership between the researcher and participant where the researcher acts as apprentice to the participant who is a master of his/her process (Holtzblatt & Beyer, 2012). CI is a method that draws significantly from ethnographic studies and can be applied as part of the task analysis stage of any software development process. In such stages, the researcher aims to uncover users' existing practices, processes, beliefs, or use of artefacts to identify opportunities to improve upon existing tasks or to specify requirements for new technology that are better grounded in user needs. This gives researchers an accurate and thorough understanding of the activity, including important details that may be overlooked when other methods are used. PD and CI can also reveal hidden elements of user’s mental models that result from the difficulty associated with verbalizing one’s process. Moreover, they can empower students to take ownership over their learning (Birch & Demmans Epp, 2015), which is atypical when other methods are used.

Current educational technology contexts reinforce existing power structures, which can contribute to adverse consequences (Avison, Baskerville, & Myers, 2001) that include the ignoring of provided analytics (Ferguson et al., 2016) or their misinterpretation and misuse (Demmans Epp & Bull, 2015). At present, CI and PD are rarely used despite their potential to inform design by better understanding learners and their environments. This potential along with a need to make learner decision-making processes explicit makes these methods crucial for designing better analytics, streamlining the design process, generating novel insights, and increasing learning analytics adoption in ways that traditional methods have not. This paper presents an exemplar study that takes these first steps.

2 CURRENT STUDY

Adult English Language Learners (ELLs) require strong writing skills to improve their work and social opportunities. In traditional classrooms, resource constraints make it a challenge to consistently provide timely and personalized feedback that support writing development. We are building a mobile application to support the writing development of mature ELLs (Age: M = 40.1, $SD = 9.2$). This study applies CI and PD guidelines to the task of designing an application that addresses the unique needs and challenges of this group of learners. We have designed and conducted a field study with 15 mature ELLs who are recent immigrants to Canada. In the first two sessions, participants completed writing samples, peer-reviewed other participants' writing samples, and participated in a one-on-one interview that explored their writing challenges and needs. The third and final session is currently underway. It consists of focus groups where ELLs actively engage in application design to generate guidelines and feature ideas for the tool through discussion, scenario-based prompts and sketching activities facilitated by the researcher. Instruments, like the Motivated Strategies for Learning Questionnaire (MSLQ), provided insight into participants' goal orientations, motivations, and beliefs. Observations of participants' writing tasks and interviews provided complementary information.

Below, we outline the advantages of CI and PD as we enacted them within this study. We explore how involving learners throughout the design process generated context-relevant insights that supplemented the results obtained through traditional methods. Some of the observations from the first two sessions are shared. We then discuss how the early incorporation of CI methods shaped the design of the focus group and the PD activities that occurred during the third session.

2.1 Advantage one: Provides context to empirical findings

Prior work stresses the importance of completing CI observations before introducing the idea of new technology (Axtell & Munteanu, 2017). This prevents participants from fixating on technology limitations or wondering how their performance will affect its design, and it helps prevent imposing pre-defined structures on analytic design. The first research objective was to study writing practices. Once these practices were understood, the next objective was to design a tool that supported natural writing flow. The CI used consisted of direct observations of participants' writing tasks and their use of help tools as well as questioning them about their workflow (when appropriate as to not disturb writing flow).

When performing peer-review activities, four ELLs were seen using their phones to translate words. When questioned, most said something similar to "I can improve my vocabulary because for the first one I think some of his words I even don't know. And I look up in my dictionary. I think it's better for me to improve the vocabulary." While ELLs understood the objective of peer-review was to provide feedback, they also viewed it as an opportunity for advancing their own learning. High MSLQ scores for both intrinsic goal orientation (M = 6, SD = 0.6 out of 7; how motivated learners are by internal factors) and task value (M = 6, SD = 0.8; importance of mastering a learning activity) were consistent with this participant claim. Observing ELLs as they performed learning activities allowed us to capture how their beliefs and motivations manifested into practice, which informed the tool's design. For instance, we could provide features to support ELL learning of unfamiliar words to expand their own knowledge as they engage in peer-review.

The interviews also revealed that many ELLs had minimal instruction on writing. For almost half, neither early schooling nor their English classes emphasized writing, as one participant shared:

> They didn't say about how to write the essay. Just our teachers said: 'You have to write three paragraphs, one paragraph about your opinion. The second paragraph it means the body, and the last paragraph you have to describe the conclusion.'

This finding was validated by an ELL instructor hired to grade the essays, who found the participants had little understanding of essay structure. In this case, combining existing practices (skill assessment) with interviews provided deeper understanding of barriers faced by ELLs, namely a systemic lack of instruction on core writing fundamentals.

2.2 Advantage two: Shines light on hidden assumptions

When asked what makes a "good" teacher, ELLs emphasized the value of praise. As one said, "It's positive. It is like the motive to continue writing because you're receiving a good feedback. Someone is praising you." Others felt unwarranted praise should not be given: "My teacher was saying all the time for me: 'oh you're doing well'. I will say: 'No, that's bad look at how many mistake' … the moment he starts saying to me 'good' that was like saying 'very bad'." This variability highlights the importance of involving learners in the design process and avoiding letting "common sense" guide design. This variability in learner personalities cannot always be captured with traditional, empirical methods. When designing technologies for real-world adoption, it is important to design for the spectrum of learners, not the average.

Beyond this, our ELLs had strong beliefs about what comprised good feedback. They had many follow-up questions on feedback they received. Their ability to articulate the feedback they wanted prompted us to reflect on our app's structure. In the initial design, writers had no direct communication with their peer-reviewer. The peer-reviewer communicated through predefined rubrics. ELLs' clarity suggested they may benefit from more direct communication with reviewers. One design to help learners access this feedback is to allow them to submit questions to guide their reviewer's assessment.

As seen here, combining results from CI with traditional assessments can provide additional insight. One major advantage of integrating both approaches is that it provides both an objective view of the learning context and the learner's perception of it. This can highlight surprising (in)consistencies between the two. Another advantage of CI is that it can help generate design ideas.

2.3 Advantage three: Identifies limitations of existing technology

Initially, a desktop app was envisioned, like most learning-to-write technologies (Schunn, Godley & DiMartino, 2016). However, interviews with the learners revealed several assumptions made by these applications that did not apply to these learners and that were not captured by the psychometric scales. First, these apps assume an instructor will manage the writing task. However, in the weeks between the first and second sessions, almost all ELLs stopped attending classes, and so, had no instructor. Second, many of these tools expect learners to compose essays, making a desktop-based application appropriate. However, most of our ELLs were job hunting or had full-time jobs and personal commitments that made regularly writing essays unfeasible. During the second session, we realized our participants required a tool that would allow them to complete short, consistent writing exercises and get feedback for improvement without instructor involvement. Thus, we began envisioning a mobile app where networks of learners provided feedback to one another on quick, daily writing exercises in a self-sustained system. Through these interviews, we constructed an understanding of the complexities in our mature ELLs' learning environments, putting us in a stronger position to start designing technology that could be integrated into learners' real-world work flow, thus addressing imbalances in their access to learning opportunities.

2.4 Advantage four: Brings theory into the real world

The final phase consists of a PD session. We drew on PD guidelines (Birch & Epp, 2015), while ensuring design decisions were supported by educational research. Participants worked in groups of three to complete sketching activities around the design of a low-fidelity user interface prototype on paper, augmented with additional props such as sticky notes. We chose this process because PD can extend theoretical principles into practical findings which can be incorporated into the development of real-world, usable technologies, provided the PD is well-grounded in theory from the start.

One important design decision for learning analytics is the information type and granularity to include (Bull & Kay, 2007). Too little information may not support a well-informed decision-making process, while too much may distract. Applying this decision in real-life contexts is challenging as it is not always clear what "too little" or "too much" looks like. One objective of our PD session is to find this balance. For instance, one feature participants design is the analytics they will receive as they complete a writing task. We have created mock-ups of several possible prompts, each of which requires different levels of learner reflection. These prompts range from short writing tips (low), to a post-writing checklist (medium), to self-assessment (high). Our goals are to have these prompts springboard design ideas that support meaningful revision without overloading the learner.

3 CONCLUSION

Though CI and PD are rarely used in educational contexts, incorporating these methods can help researchers gain a more holistic understanding of learners and the learning context, as illustrated by our study. We found the psychometric scales, synthesized with CI methods, helped provide a comprehensive and holistic understanding of both the challenges and needs of ELLs learning to write. CI and PD complement existing practices, most importantly, they can guide researchers in drawing theory into practice. Preliminary analyses, including the designs generated from PD, challenges and suggestions for future directions will be discussed at the workshop.

REFERENCES

Avison, D. E., Baskerville, R., & Myers, M. (2001). Controlling Action Research Projects. *Information Technology & People*, *14*(1), 28–45.

Axtell, B. & Munteanu, C. (2017). Using frame of mind: documenting reminiscence through unstructured digital picture interaction. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, Article 50, 4 pages.

Birch, H., & Demmans Epp, C. (2015). Participatory design with music students: Empowering children to develop instructional technology. Presented at the American Educational Research Association (AERA) Annual Meeting, Chicago, Illinois.

Bull, S., & Kay, J. (2007). Student Models that Invite the Learner in: the SMILI:) Open Learner Modelling Framework. *International Journal of Artificial Intelligence in Education (IJAIED)*, *17*(2), 89–120.

Demmans Epp, C., & Bull, S. (2015). Uncertainty representation in visualizations of learning analytics for learners: Current approaches and opportunities. *IEEE Transactions on Learning Technologies*, *8*(3), 242–260. https://doi.org/10.1109/TLT.2015.2411604

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., … Ullman, T. (2016). *Research Evidence on the Use of Learning Analytics - Implications for Education Policy* (Joint Research Centre Science for Policy Report No. EUR 28294 EN). European Commission.

Holtzblatt, K., & Beyer, H. R. (2012). 8.1 8.1 Motivations and Key Principles. The Encyclopedia of Human-Computer Interaction, 2nd Ed.

Schunn, C. D., Godley, A. J., & DiMartino, S. (2016). The reliability and validity of peer review of writing in high school AP English classes. Journal of Adolescent & Adult Literacy.

Muller, M. J., & Kuhn, S. (1993). Participatory design. Communications of the ACM, 36(6), 24-28.

Muller, M. J. (2003). Participatory design: the third space in HCI. Human-computer interaction: Development process, 4235, 165-185.

Wixon, D., Holtzblatt, K., & Knox, S. (1990). Contextual design: an emergent view of system design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 329–336.

Workshop: Methodological Bases for the Measurement of Learning in Learning Analytics

Sandra Milligan[1], Linda Corrin[2], Nancy Law[3], Ulla Ringtved[4]
[1]Melbourne Graduate School of Education, University of Melbourne, Australia, s.milligan@unimelb.edu.au
[2]Williams Centre for Learning Advancement, University of Melbourne, Australia, linda.corrin@unimelb.edu.au
[3]Faculty of Education, University of Hong Kong, Hong Kong, nlaw@hku.hk
[4]University College of Northern Denmark, Denmark, ulr@ucn.dk

ABSTRACT: This Workshop focuses on the methodology of learning analytics. It aims to promote communication between two communities of scholars – learning analysts and educational measurement specialists. The argument is that strength will accrue from methodological collaboration across the fields, which share an interest in learning, a commitment to improving practice and a belief in the power of analysis. They may differ in how the construct of learning is understood, and what is understood by the term 'measured'. Different criteria may be applied when assessing quality of data, and the standards of proof required as to the utility and interpretability of outcomes. Different data modeling techniques are used to uncover meaning in data. This workshop will provide opportunities for expert methodologists from both fields to collaborate, in the company of representatives of key stakeholders such policy makers and public officials, in the interests of improving trustworthiness, validity, reliability, utility and interpretability of analytics used in assessment and measurement of learning. A Workshop Report will summarise the opportunities for, and likely outcomes of improve collaboration between the fields, and if warranted, the organisers will lead an initiative for the establishment within SOLAR of an ongoing Special Interest Group on Measurement Analytics.

Keywords: Measurement modeling, methodology, learning analytics, educational measurement, assessment of learning, learning, measurement analytics

1 BACKGROUND:

This Workshop focuses on the methodology of learning analytics, aiming to extend the conversation between two communities of scholars – learning analysts and educational measurement specialists – to the benefit of both. The fields share an interest in learning, a commitment to improving practice and a belief that data can assist understanding of learning. Both fields have an interest in measuring learning. There are also differences, and these provide opportunities for productive collaboration.

The methodology of learning analytics is concerned principally with interrogation and interpretation of digital data harvested from digital educational applications such as LMS platforms or games, or from data collection devices including wearable, audio or video recordings or other data capture devices embedded in the environment. The plethora of learning-related information charts social interactions, eye-gaze direction, facial expression, and a range of other physical, physiological emotional or neurological indicators. The *process* of learning is traced as well as the outcomes. Learning analysts applies techniques such as social network analyses, data mining, machine learning, semantic analysis and so on.

The field of learning analytics is young and is not without it challenges. There is growing awareness that measures of learning need to be accurate, fair, reliable, valid, and interpretable regardless of whether they are used for prediction, for feedback, or for research (Berger et al. 2017; Milligan, 2015; Prinsloo & Slade, 2017; Ringtved, Milligan, Corrin & Law, 2017). The Leaning Analytics Community Exchange (LACE) recently registered concerns about use of big data in education: data can do harm if used to shape the information or treatment a person, if based on faulty inferences, especially if decisions are made on the basis of automated algorithms. Questions are being raised about the effectiveness of analytics (Ferguson & Clow, 2017). Inferring attributes from a characterisation of statistical categories is insufficient to engender trust in the patterns "found" in raw data.

Educational measurement also has at its core the analysis of large scale quantitative data on learning, but this field is older, and is concerned principally to use data to derive assessments of human attributes that are reliable, valid, have utility, and are interpretable for educationalists (Messick, 1995, Wilson, 2005), It especially focuses on measuring learning-related attributes of learners i.e., what learners know or can do. There is a well-established methodology, underpinned by understandings that data cannot speak for itself, that every relationship found in data is not meaningful, and that some are damaging if used to predict or shape learning. Educational measurement techniques provide a means to cut through the inherent complexity and interrelatedness of educational evidence to distinguish what is meaningful and useful, from what is merely related.

Although not young, educational management is not without its challenges either. Its job of is getting harder. Changes in conceptions of *what* learning should be assessed are evident in reforms of national and international curriculum frameworks, which now routinely supplement the cognitive outcomes of traditional subjects and disciplines with requirements that learners develop complex competencies comprised of knowledge, values, attitudes, skills and beliefs required for effective performance in any field. These traits are difficult to assess using traditional approaches and traditional data forms. Teaching methods are changing too. Digital learning platforms and applications classes are ubiquitous. Greater reliance is placed on automated assessments, and agents. Educational measurement and assessment is increasingly using big data of the kind that learning analysts engage with, and its models, techniques and tools are needing to change at the same time (Mislevy, 2016, Pellegrino, 1999).

The advantages of methodological collaboration between these two fields have been remarked in both the learning analytics community, and the educational measurement community (Drasgow, 2016; He et al., 2016). There are advantages in exploring differences between the fields in assumptions about the nature of learning and how learning can be indicated and understood, even in what is understood by the term 'measured'. Different assumptions may apply to consideration of matters of data adequacy, and control, and the standards of proof required as to the utility and interpretability of findings. The fields use different statistical techniques for data modeling, and for uncovering meaning in the data. There is, however, already evidence that collaboration between the two fields can prove productive, including the emergence of teams combining methodologies to good effect (Milligan, 2015; Griffin & Care, 2015; Shute & Ventura 2009)

2. PURPOSE OF THE WORKSHOP

In this context, the aim of this workshop is to extend the methodological collaboration between the learning analytics community and the educational measurement community, by convening a group of methodology-focused researchers, and other key stakeholders interested in the measurement of learning, to discuss and assist productive collaboration.

If the discussion warrants, the organisers will present an argument for the establishment of an ongoing Special Interest Group on Measurement Analytics, within SOLAR, aimed at stimulating methodological collaboration within the learning analytic and measurement communities. Organisers will seek to engage participation with measurement-focussed organisations such as the National Council for Measurement in Education.

3. WORKSHOP ORGANISATION

A full day workshop is expected to attract about 50 participants. Participation is sought from a range of interest groups, including, inter alia: DesignLAK16 and DesignLAK 17 participants; learning analytics researchers and practitioners; ASCILITE learning analytics and e-assessment SIGs; and the National Council of Educational Measurement. It is also expected to involve a number of advisors to policy makers and public officials with an interest in the validity, reliability, utility and interpretability of analytics for assessment and measurement in education. The workshop space is arranged round tables seating approximately 6 people. Equipment includes butchers paper and pens on each table, a lectern and data projection equipment that manages BYOG devices. Wifi is required.

1.1 Pre-workshop planning

A Workshop website will facilitate discussion and interaction of the developing community. A twitter hash tag and mailing list will be established to facilitate communication. A call for abstracts of 400 – 500 words explaining methodology and showcasing the methodological feature of work. will be directed to invited expert methodologists working on the measurement of learning, within or across the two fields. The call will also be open to the general LAK community. The workshop organisers will review submissions, leading to selection of up to 8-10 case studies of methodological approaches. Organisers will also invite a panel of discussants expert in methodology in each of the fields of learning analytics and/or educational measurement.

1.2 The workshop design

The bulk of the day will be organised around three main working sessions, each comprising three different elements and employing the technique of World Café[1] to facilitate knowledge building and networking. Each working session will include two or three of the presenters explaining the methodological approach, and the working principles about learning that lie behind it. The invited expert discussants will provide commentary on the presentations, teasing out opportunities for building on the perspectives of each field. They are likely to focus on: assumptions about the nature of learning and how learning can be indicated and understood; the purposes of measurement and what is understood by the term 'measured'; the standards adopted in relation to data adequacy and control; assumptions about what constitutes proof of utility and interpretability of findings; the means used to uncover meaning in the data; and the appropriateness of data modeling approaches. All participants will then actively engage in knowledge building, collaboratively synthesising a set of 'best practice' methodological principles derived from the presentations and discussant inputs.

1.3 Outcomes

A range of workshop outcomes is envisaged. First, presenters will be invited to publish their presentation abstracts on the workshop website in the weeks before the workshop. Second, presenters will be encouraged to present their finalised papers in the Companion Proceedings. Third, a Workshop Report and paper will be prepared by the workshop organisers, summarising the opportunities and likely outcomes of improve collaboration between the fields of learning analytics and educational measurement. Fourth if warranted, and to maintain momentum, the organisers will develop a proposal for SOLAR to establish a SIG in the area.

REFERENCES

Buckingham Shum, S. and Crick, R. D. (2016).Learning Analytics for 21st Century Competencies. Journal of Learning Analytics, 3, 3

Draschler, H. and Greller, W. (2016) Privacy and Analystics – its a DELICAT issue. A Checklist for Trusted Learning Analytics. In Proceedings of the LAK '16 (Edinburgh, United Kingdom, April25-29.

Drasgow, F. (2016) Technology and Testing: Improving Educational and Psychological Measurement, Taylor and Francis.

Fergurson, R. and Clow, D. (2017). Where is the evidence? A call to action for learning analytics. In proceedings of the Learning Analytics and Knowledge Conference, Simon Frazer University , Vancouver, British Columbia, Canada.

Griffin, P. & Care, E. (2015) Assessment and Teaching of 21st Century Skills: Methods and Approaches. Springer, 2015.

[1] http://www.theworldcafe.com/key-concepts-resources/world-cafe-method/

He, J., Rubinstein, B. I. P., Bailey, J., Zhang, R., Milligan, S., & Chan, J. (2016). MOOCs meet measurement theory: A topic-modelling approach. Paper presented at the AAAI Conference, Phoenix, Arizona.

Milligan, S. K. (2015). Crowd-sourced learning in MOOCs: learning analytics meets measurement theory. In Proceedings of the Learning Analytics and Knowledge Conference Poughkeepsie, NY, USA.

Milligan, S. K. & Griffin, P. (2016) Understanding learning and learning design in MOOCs: a measurement-based interpretation. Journal of Learning Analytics Special Section on Learning Analytics for 21st Century Competencies,3, 2.

Mislevy R.J. A Dispatch from the Psychometric Front, In Proceedings of the Learning Analytics and Knowledge Conference (Edinburgh, Scotland, UK, 2016).

Pellegrino, J.W. (1999) The evolution of Educational Assessment: Considering the past and imagining the Future. Willian H Angoff memorial Lecture, ETS

Prinsloo, P.; and Slade, S. (2017) An elephant in the room. In proceedings of the Learning Analytics and Knowledge Conference, Simon Frazer University, Vancouver, British Columbia, Canada.

Ringtved, U.; Milligan, S.K.; Linda Corrin, L.& Nancy Law, N. (2017). DesignLAK17: Quality metrics and indicators for analytics of assessment design at scale, 7th Annual Learning Analytics and Knowledge Conference, Simon Frazer University , Vancouver, British Columbia, Canada.

Shute, V. & Ventura (2013). M. Stelth Assessment:Measuring and Supporting Learning in Video

Wilson, M. (2005) Constructing Measures: An item Response Modeling Approach. Taylor & Francis Group, New York.

Measuring Individual Learning Progress the Combinatorial Way

Michael D. Kickmeier-Rust
University of Teacher Education, St. Gallen, Switzerland
and
University of Technology, Graz, Austria
michael.kickmeier@phsg.ch

ABSTRACT: Measuring human aptitude and learning processes has a very long tradition and the research community established well-elaborated test theories. Since the advent of learning analytics, a new community is contributing to the theories and methodologies of measuring learning. In this paper we introduce Competence-based Knowledge Space Theory as a combinatorial, multi-dimensional framework for modelling and assessing competencies and competence development over time. A particular strength of the approach is the conceptual and methodological separation of latent aptitude (competencies, knowledge) and observable performance as indicators for aptitude. In the Lea´s Box project, the approach has been realized and deployed as online learning analytics platform tailored to the needs of conventional classroom settings (i.e., only little, heterogeneous, and incomplete data as basis for analytics). The paper illustrates the potential strengths of the approach and gives and outlook for future developments.

Keywords: Psychometrics, Competence-based Knowledge Space Theory, Competence Modelling, Learning Analytics

1 MEASURING APTITUDE

Abilities, strength, weaknesses of learners, their knowledge and misconceptions, their needs and goals are extremely diverse and rich. In essence, these are the core determinants of learning and performance, school success, and – in the end – planning individual teaching and support. To realize a successful approach to teaching we need methods and techniques to support teachers in assessing learning/development progress and abilities (knowledge, competencies, or skills) in a formative way and to provide the learners with appropriate and tailored support and guidance.

Despite a considerable trend towards such formative assessment and competence-oriented schooling, still, in the majority of classrooms, evaluating learning performance and achievements is reduced to gauging achievements and knowledge with a single numerical value – the school grades. Most often, these summaries are based on superficial, one-dimensional tests and test items. This approach, however, cannot express what learners really know or are able to; characterizing proficiency by a single variable at best suffice basic fail/pass decisions, as argued by Mislevy and colleagues (2003). A good example for the weakness of the approach is the I.Q. (intelligence quotient), which attempts to characterize all the various abilities, strength and weaknesses of a person in many categories and disciplines (math, language, cognition, memory, etc.) with a single numerical value, in the end. The origin of this popular test theoretical approach lies in 19th century physics and the occurrence of disciplines like “anthropometry”, the “art of measuring the physical and mental faculties of human beings”. Prominent proponents were Francis Galton, William Kelvin, or Carl Pearson. The predominant tenor was, if you cannot measure it, it is not science. Kelvin, for example, said “If you can‘t assign an exact numerical value, express it in numbers, your knowledge is of a meagre and unsatisfactory kind” (cf. Falmagne, Cosyn, Doignon, & Thiery, 2003).

When aiming at an evidence-centered and formative approach to evaluating achievements and proficiency in school settings, when focusing on a formative approach to appraisal with the idea of supporting learners in a meaningful way and on an individual level, a deeper and more precise understanding is necessary. Such attempt, however, is not trivial. It is complex, demanding, and costly. On the one hand it takes a profound theoretical approach to evaluation which includes all the various dimensions and on the other hand, evaluators (teacher, in the first instance) are required to develop a fair and comprehensive image of learners, their origin, learning performance and present ability/proficiency/knowledge for each individual learner. Non-numerical test theories provide ideas for such broadened evaluation. Basically with the rise of mathematical combinatorics and with the rise of powerful computer technologies the more demanding approaches to precisely describe the various abilities of a learner along multiple dimensions appear more feasible in practice. More formative and competence-centered learning analytics should focus on two fundamental concepts; that of the so-called substantive features, which concern the characteristics of the learning domain and the learning process, and the evidentiary-reasoning aspect, which concerns the information we can draw from the learners’ behaviors. It takes a formal framework that links both, the substantive and the evidentiary-reasoning aspects of a sound, reliable, and, in a way, formative assessment. Most likely, such frameworks are based on probability values, for example, Item Response Theory (Van der Linden, & Hambleton, 1997), Latent Class Models (Collins, & Lanza, 2010), or Bayesian inference networks (Jensen, 1996).

In this paper, I intend to present a structural approach to learning analytics based on Knowledge Space Theory (KST), founded by Doignon and Falmagne (1985, 1999) and extensions such as Competence-based Knowledge Space Theory (CbKST) that melds both, the substantive and the evidentiary-reasoning, coming from the genre of autonomous intelligent and adaptive tutoring systems. The idea was to broaden the ideas of the linear Item Response Theory (IRT) scaling, where a number of items are arranged on a single, linear dimension of “difficulty”. In essence, KST provided a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. More recent advancements of the theory accounted for a probabilistic view of test results and they introduced a separation of observable performance and the actually underlying abilities and knowledge of a person. Such developments lead to a variety of theoretical, competence-based approaches (see Albert & Lukas, 1999 for an overview). An empirically well-validated approach to CbKST was introduced by Korossy (1999); basically, the idea was to assume a finite set of more or less atomic competencies (in the sense of some well-defined, small scale descriptions of some sort of aptitude, ability, knowledge, or skill) and a prerequisite relation between those competences.

2 COMBINATORIAL COMPETENCE MODELLING

Combinatorics is the area of mathematics that is concerned with the enumeration (counting) of specified structures, the existence of such structures that satisfy certain given criteria, the construction of these structures, perhaps in many ways, and optimization, as finding the "best" structure or solution among several possibilities, be it the "largest", "smallest" or satisfying some other optimality criterion. This definition appears very suitable and sensible for being used to understand and model the concept of human aptitude (competences, skills, knowledge, etc). A visual ad intuitive approach to tackle a combinatorial understand of competencies is *Hasse diagrams*.

A Hasse diagram (cf. Skiena, 1990) is a strict mathematical representation of a so-called semi-order in form of a directed graph that reads from bottom to top. A semi-order is a type of mathematical ordering of a set of items with numerical values by identifying two items as equal or comparable if the values are within a given interval of error or noise. Semi-orders were introduced in mathematical psychology by Duncan Luce in 1956 in human decision research without the assumption that indifference is transitive. This approach is also crucial for handling human learning and the resulting performance that is prone to all sorts of errors and peripheral aspects (perhaps failing in a test although the learner holds the knowledge due to being tired). A Hasse diagram is one way of displaying such ordering – in our case competences or competency states (which is to be explained in the following section). The technique was invented in the 60s of the last century by Helmut Hasse. The diagram exists of entities (the nodes), which are connected by relationships (indicated by edges).

The mathematical properties of a semi-order and the Hasse diagrams are (i) reflexivity, (ii) anti-symmetry, and (iii) transitivity. Reflexivity refers to the view that an item, perhaps a competency, references itself in a cause/effect sense. Anti-symmetry demands that if one entity is a prerequisite of another, this relationship is not invertible; as an example, if competency x is a prerequisite to develop competency y, y cannot be the perquisite of competency x. Finally, transitivity means that whenever an element x is related to an element y, and y is in turn related to an element z, then x is also related to z. In principle, the direction of a graph is given by arrows of the edges; by convention however, the representation is simplified by avoiding the arrow heads, whereby the direction reads from bottom to top. In addition, the arrows from one element to itself (reflexivity property), as well as all arrows indicating transitivity are not shown in Hasse diagrams. The following image (Figure 1) illustrates such a diagram. Hasse diagrams enable a complete view to (often huge) structures. Insofar, they appear to be ideal for capturing the large competence or learning spaces occurring in the context of assessment and learning recommendations (for example, all the competencies involved in the math curriculum for a specific age).

In an educational context, a Hasse diagram can display the non-linear path through a learning domain starting from an origin at the beginning of an educational episode (which may be a single school lesson but could also be the entire semester). Moreover, the elements in the diagram may refer to (latent) competencies, to learning objects or test items. Figure 1 illustrates the simple example of typical learning objects in a certain domain. The beginning of a learning episode is usually shown as { } (the empty set) at the bottom of the diagram. Now a learner might attend two learning objects (v and x), which is indicated by the edges; this, in essence, establishes three possible learning paths. After x, as an example, this learner might attend w, or v but not u yet, which in turn opens further three branches for the learning path until reaching the final state, within which all learning objects have been attended.

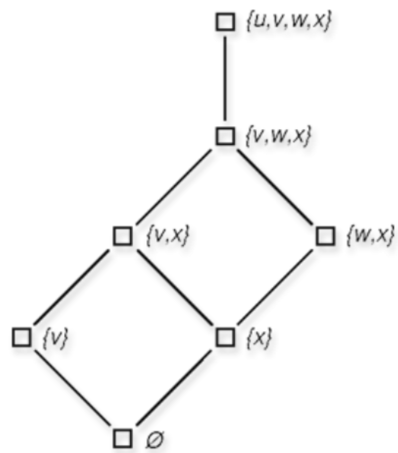


Figure 1: An exemplary Hasse diagram

As claimed initially, in the context of formative LA, a competence-oriented approach is necessary. Thus, a Hasse diagram can be used to identify and display the latent competencies of a learner in the form of so-called competence states. An elaborated theoretical approach to do so is Competence-based Knowledge Space Theory (CbKST). The approach originates from Doignon and Falmagne (1985, 1999) and is a mathematical psychological, set-theoretic framework for addressing the relations among problems (e.g., test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. While the original Knowledge Space Theory focuses only on performance (the behavior; for example, solving a test item), its extension CbKST introduces a separation of observable performance and latent, unobservable competencies, which determine the performance (albert & Lukas, 1999). This is a psychological learning-theoretical approach, which highlights that competencies (e.g., the ability to add two integers) are unobservable latent constructs and which can only be observed or assessed indirectly.

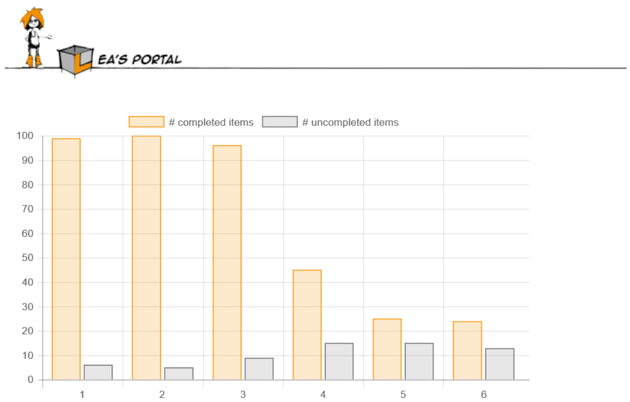
We interpret the performance of a learner (e.g., mastering an addition task) in terms of holding or not holding the respective competency. In addition, recent developments of the approach are based on a probabilistic view of having or lacking certain competencies. In our example, mastering one specific addition task allows the conclusion that the person is able to add two numbers (to hold this competency) only to a certain degree or probability. When thinking of a multiple-choice item with two alternatives, as another example, mastering this item allows only to 50 percent that the person has the required competencies/ knowledge. On the basis of these fundamental views, CbKST is looking for the involved entities of aptitude (the competencies) and a natural structure, a natural course of learning in a given domain. For example, it is reasonable to start with the basics (e.g., the competency to add numbers) and increasingly advance in the learning domain (to subtraction, multiplication, division, etc.). As indicated above, this natural course is not necessary linear, which bears significant advantages over other learning and test theories. As a result we have a set of competencies in a domain and potential relationships between them. In terms of learning, the relationships define the course of learning and thus which competencies are learned before others. In CbKST such relationships are called prerequisite relations or precedence relations. On the basis of competencies and relationships, in a next step, we can obtain a so-called competence space, the ordered set of all meaningful competence states a learner can be in. As an example, a learner might have none of the competencies, or might be able to add and subtract numbers; other states, in turn, are not included in this space, for example it is not reasonable to assume that a learner holds the competency to multiply numbers but not to add them. By the logic of CbKST, each learner is, with certain likelihood, in one of the competence states.

3 STRUCTURAL, THORY-DRIVEN LEARNING ANALYTICS

Recent advancements of CbKST primarily concern the integration of theories of human problem solving (given that most indictors can be interpreted as solving some sort of problem). This work was essentially driven in the genre of smart, educationally adaptive computer games for learning – loosely speaking for developing an educational AI support the players of the game (Kickmeier-Rust et al., 2010, Kickmeier-Rust & Albert, 2010) Also, approaches to communicate and visualize the results of analytics in form of Open Learner Models have been elaborated (Ginon et al., 2016).

In the European project Lea's Box (www.leas-box.eu), moreover, the CbKST approach has been implemented in a comprehensive learning analytics platform. The focus of this learning analytics toolbox (Lea's box) is on supporting teachers in their concrete school settings with suitable learning analytics features. A main challenge, for example, is that typical school scenarios are characterized by the lack of a reasonable and coherent basis of data. Specifically K12 / K18 education is still an analogous process and data a generated sparsely and usually they are of a very heterogeneous nature. For example, teacher may use a variety of different learning apps or games, homework may be done using Google docs, or little tests and quizzes may be presented in an e-Learning platform. The big challenge is to access all these sources, to aggregate the data, and to make holistic analyses on their basis. The CbKST approach enables this aggregation by its competence-performance separation, which means that various data sources serve as evidence for certain competences or skills. In other words, Learning Analytics usually requires a robust and possibly homogenous set of digital data as basis for analyses. In typical classroom settings, this is simply not the case. If not big than there is little data I the classroom. Teachers are using learning management systems on a lose basis, perhaps they use learning apps or serious games every now and then, perhaps the give homework with Google docs, and usually they are required to keep records about student achievements. By establishing a central latent competence model, all those diverse and incomplete sources can be utilized as evidence for the competencies a student holds – with a certain cautious probability. Thus, with each little information and with each little achievement, the believe model is getting clearer, more stable, and more valid. In the Lea´s Box project, we developed a prototypical Learning Analytics platform on the theoretical basis of CbKST (and other similar mathematical-psychological approaches). Figures 2 and 3 provides two screen captures of the online system. The upper shows conventional performance statistics of students in terms of completed versus uncompleted tasks.

The lower image shows the teacher view of a Hasse diagram; the left panel is a snapshot of a realistic large competence structure for the domain Mathematics. The bullets indicate the possible competence states. The probabilities that a learner is in a particular competence state are color coded – the darker and more blue, the higher is the probability. The system is capable to display the most likely learning path of learners, that is, the history of how this believe model unfolded over time. The dotted line is a prediction of future learning steps (cf. Kickmeier-Rust, in press).



Learning Object	Completed Items	Uncompleted Items	Avg. Time	SD of Avg. Time
1	99	6	936.72	466.93
2	100	5	311.93	247.27
3	96	9	267.89	90.46
4	45	15	1075.38	648.73

Figure 2: Screen capture of the Lea´s Box system

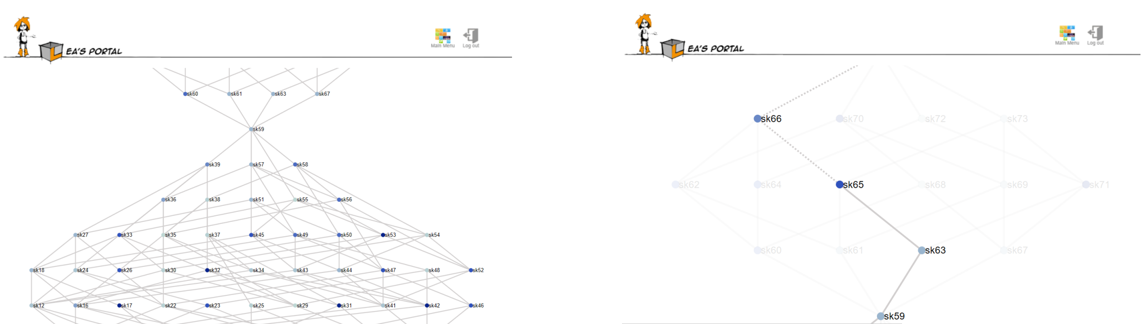


Figure 3: Hasse diagram visualization displaying a competence model (left) and an individual learning paths (right)

4 CONCLUSIONS

There is little doubt that frameworks, techniques, and tools for learning analytics will increasingly be part of a teacher's professional life in the near future. However, a major critique is that conventional approaches focus too much on a summative measurement, without given formative, actionable information to teachers. In addition, learning analytics is too often a bottom-up data-driven undertaking, which in turn lacks advice for the next best steps for each individual learner. The presented approach may be contribution to the learning analytics community, helping to make conventional teacher in conventional classroom settings benefit from the most recent technologies. In terms of psychometrics, the CbKST approach may serve a promising complement to the rather one-dimensional psychometric theories, such as IRT, by adding a large spectrum of diversity to the measurement model. Pilot studies with the Lea´s Box system clearly indicate that teachers very much appreciate the information that can be obtained by the approach. However, equally clear is that thy typical visualization methods – the Hasse diagrams – are not suitable to be easily understood by teachers. Future research activities must develop more intuitive and simpler forms of visualizing the information the inherently is held by the Hasse diagrams.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7) under grant agreement no. 619762 (LEA´s BOX). This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

REFERENCES

Albert, D., & Lukas, J. (Eds.). (1999). Knowledge spaces: Theories, empirical research, and applications. Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, L.M., & Lanza, S.T. (2010). Latent class and latent transition analysis for the social, behavioral, and health sciences. New York: Wiley.

Doignon, J.-P., & Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. International Journal of Man-Machine Studies, 23, 175-196.

Doignon, J.-P., & Falmagne, J-C. (1999). Knowledge spaces. Berlin: Springer.

Falmagne, J.-C., Cosyn, E., Doignon, J.-P., & Thiery, N. (2003). The assessment of knowledge, in theory and in practice. In R. Missaoui & J. Schmid (Eds.), Lecture Notes in Computer Science, (pp. 61–79). Berlin: Springer.

Ginon, B., Johnson, M D., Turker, A., Kickmeier-Rust, M. D. (2016). Helping Teachers to Help Students by using an Open Learner Model. In R. Vatrapu, M. D. Kickmeier-Rust, B. Ginon, & S. Bull (Eds.), Proceedings of the Fourth International Workshop on Teaching Analytics, in conjunction with EC-TEL 2016 (pp. 23-29). September 16, 2016, Lyon, France.

Kickmeier-Rust, M. D. (in press). The Learning Horizon: Predicting Individual Learning Performance on the Basis of Formal Competence Models.

Kickmeier-Rust, M. D., & Albert, D. (2010). Micro adaptivity: Protecting immersion in didactically adaptive digital educational games. Journal of Computer Assisted Learning, 26, 95-105.

Kickmeier-Rust, M.D., Mattheiss, E., Steiner, C.M., & Albert, D. (2011). A Psycho-pedagogical framework for multi-adaptive educational games. International Journal of Game-Based Learning, 1 (1). 45-58.

Korossy, K. (1999). Modelling knowledge as competence and performance. In D. Albert & J. Lukas (Eds.), Knowledge spaces: Theories, empirical research, and applications (pp. 103–132). Mahwah, NJ: Lawrence Erlbaum Associates.

Jensen, F.V. (1996). An Introduction to Bayesian Networks. Berlin: Springer.

Luce, R. D. 1956. Semiorders and a theory of utility discrimination. Econometric,a 24, 178–191.

Mislevy, R.J., Steinberg, L.S., & Almond, R.G. (2003). On the structure of educational assessments. Measurement: Interdisciplinary Research and Perspectives, 1, 3-67.

Skiena, S. "Hasse Diagrams." §5.4.2 in Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica. Reading, MA: Addison-Wesley, p. 163, 169-170, and 206-208, 1990.

Van der Linden, W.J., & Hambleton, R.K. (Eds.) (1997). Handbook of modern item response theory. New York: Springer.

Exploring the Relationship between Student's Emotional Factors and Achievement in a MOOC Course

Shuai Zhao
China Institute for Educational Finance Research
Peking University, Beijing, China
zhaoshuai2013@pku.edu.cn

Xiaoting Huang
China Institute for Educational Finance Research
Peking University, Beijing, China
xthuang@pku.edu.cn

ABSTRACT: Researchers have examined various behavioral variables that might affect student achievements regarding massive open online courses (MOOC). However, little attention has been paid to student's emotional factors. In this paper, we detected students' positive and negative sentiments from forum posts of a Chinese MOOC course using a lexicon-based method, and investigated the relationship between the sentiments and students' final grades. After controlling for behavioral variables such as times of viewing lecture videos and reading forum posts, we found that positive sentiment was not significantly correlated with student's final grade, whereas negative sentiment had moderately strong correlation with grade. One plausible explanation was that negative sentiment detected in this case did not reflect student's attitude towards the course, instead, it showed student's achievement anxiety as they needed assistance to complete their assignments or quizzes. Under most circumstances in this study, these students received help from TAs and their classmates, which may, in turn, improve their grades. In summary, our findings suggested that using forum data to detect student's need for help and addressing their needs would help promote online learning outcomes.

Keywords: MOOC, Emotional factors, Achievement, Text analysis

1 INTRODUCTION

MOOC has attracted much attention worldwide as it promises to make educational resources easily accessible to all. But in practice, the dropout rates of MOOC reached as high as 90% (Koller, 2012). Hence, the effect of MOOC is put in question. How to improve course designs to encourage more students to complete the online courses becomes a key issue. Researchers have examined various behavioral variables that might influence student achievement. For instance, Jiang, Zhang and Xu (Zhuoxuan, Yan, & Xiaoming, 2015) found that times of viewing lecture videos and times of submitting quizzes had some power in predicting the completion rate. Yang et al. (Yang, Sinha, Adamson, & Rosé, 2013) showed that posting behaviors on the forum such as number of posts were related to the dropout rate. Ramesh et al. (Ramesh, Goldwasser, Huang, Daumé III, & Getoor, 2013) concluded that engagement type was a significant predictor of students' achievements.

However, little attention has been given to emotional factors, which were found to influence student achievements in traditional classrooms. For instance, Pritchard and Wilson (Pritchard & Wilson, 2003) showed that students' anxiety negatively correlated with GPA. Villavicencio and Bernardo (Villavicencio & Bernardo, 2013) recorded that positive emotions could enhance the positive correlations between self-regulation and academic achievement. Mega and Beni (Mega, Ronconi, & De Beni, 2014) claimed that positive (negative) emotions indirectly improved (lowered) academic achievement by influencing learning motivation. But not all the existing studies are in consensus about the relationship between sentiments and achievements. Chen's research (Q. Chen, 2007) indicated that there was no significant correlation between student emotions and GPA.

In this paper, we explored the relationship between student sentiments and achievements in a Chinese MOOC course. We hoped to find out whether emotional factors may play a more important role in self-driven online courses. In addition, we used forum data to extract sentiment scores instead of traditional questionnaires. Using this approach might give us more insight on how emotional factor work during the process of the course, and in turn provide suggestions on how to better facilitate learning online.

2 METHOD

We analyzed the data of **Introduction to Computation A**, a course opened by Peking University on Coursera platform in this study. Some basic information of the course is listed in Table 1.

Table 1: Some basic information of the course.

Start Date	End Date	Registrants	Dropout Rate	Pass Rate	Distinction Rate
Sep, 2014	Jan, 2015	14,855	90.9%	1.0%	0.5%

A student who passed the course gained a final grade higher than 60.
A student who achieved the “distinction” level gained a final grade higher than 85.

To examine the relationship between sentiments and grade, we first extracted positive and negative sentiments from all of the posts in the class forum using a lexicon-based method (R. Chen & Lazer, 2013; Zhao, Tong, Liu, & Tan, 2016). Compared with questionnaire-based methods, the advantages of lexicon-based methods lie in two aspects: 1) they overcome the shortcomings of self-report questionnaires. As Villavicencio and Bernardo (Villavicencio & Bernardo, 2013) stated, although self-report measures for sentiment evaluation could be well-established, they were still not ideal since sometimes it was not easy for students to assess their own emotions accurately; 2) they capture the sentiments of students over the course duration, instead of taking a snapshot as most questionnaires usually do.

There are totally 450 forum discussants among all the students. For each one of them, we aggregated all of his/her posts to form a document. Table 2 gives some descriptive statistics of their numbers of posts and document lengths[1].

Table 2: Some descriptive statistics of forum discussants.

	Posting Times	Document Length
Maximum	301	25,239
Minimum	1	3
Average	9	826
Total	3,864	526,083

We extracted positive and negative sentiments for each forum discussant from his/her document. There are two kinds of methods that are commonly used to collect sentiment information from texts. One is machine learning methods, whose procedures mainly consist of two steps. First, train classifiers on texts with sentiment labels. Then, use the classifiers to determine new texts’ sentiments. The other one is lexicon-based methods, whose key idea is to identify sentiment words from texts according to a predefined sentiment dictionary. Note that machine learning methods can only recognize texts’ sentiment polarity, but not their polarity strength. Hence, we chose lexicon-based methods in this study.

[1]The unit of a document is a character.

The sentiment lexicon we used is augmented NTU sentiment dictionary (ANTUSD) (Wang & Ku, 2016), which covers six large Chinese sentiment corpora from year 2006 to year 2010, and contains 9,527 positive words and 11,278 negative words. ANTUSD gives each sentiment word its polarity strength. The larger the absolute value of the strength[2], the stronger the sentiment expressed by corresponding word. The sentiment information in ANTUSD is quite valid, as Wand and Ku (Wang and Ku, 2016) has shown that based on ANTUSD the *F*-score of sentiment polarity classification reached 98.21%. Besides ANTUSD, we found some unique sentimental expressions on the forum. In order to capture the sentiments comprehensively, we manually picked 66 sentiment words and icons from the posts, consisting of 31 positive ones and 35 negative ones. We then appended those words and icons to ANTUSD, and assigned them sentiment polarity strength based on their synonyms in ANTUSD. Table 3 lists some of the sentiment icons as well as their related information.

Table 3: Some extracted sentiment icons and their related information.

Sentiment Icons	Synonyms	Polarity Strength
ORZ	worship	0.021
O[∩_∩]O	happy	0.434
^_^	happy	0.434
:-D	happy	0.434
2333	laugh	0.282

For each forum discussant, we first segmented his/her document, then deleted stop words, and finally calculated his/her positive and negative sentiment scores using Equations 1 and 2 respectively.

$$Poscore = \frac{\sum_{w \in P_W} s_w \times f_w}{N} \tag{1}$$

$$Negscore = \frac{-\sum_{w \in N_W} s_w \times f_w}{N} \tag{2}$$

In the equations, P_w and N_w separately represent positive word set and negative word set that appear both in the document and the updated ANTUSD, and word w is one of them. s_w and f_w stand for sentiment polarity strength and frequency of w respectively. N is the total number of words in the document. Note that *Negscore* is greater than 0 according to Equation 2. Larger

[2]The strength value for a positive (negative) word is positive (negative).

Negscore indicates stronger negative sentiment. All the document analyses and calculations are based on "*jieba*" package in Python 2.7.13 (Embedded in the environment of Anaconda2 4.4.0, which is available at https:// repo.continuum.io/archive/). Some forum discussants' documents are too short and contain only stop words. We delete those students' data and finally obtained the records of 436 forum discussants.

After we had the composite sentiment scores, we built a regression model to examine the relationship between the sentiments and student grades. In the model, we controlled some well acknowledged behavioral variables that can influence achievements, including the times of viewing lecture videos and reading forum posts. We list all the variables and their short descriptions in Table 4.

Table 4: Variable descriptions.

Variables	Descriptions
TLv	Lecture viewing times
TLd	Lecture downloading times
TFr	Forum reading times
TFp	Forum posting times
Poscore	Positive sentiment score
Negscore	Negative sentiment score
Grade	Final grade

3 RESULTS AND ANALYSIS

As the majority of MOOC courses, most students of **Introduction to Computation A** did not provide their demographic data. Hence, we did not implement achievement analyses based on variables like age and gender. Our key task is to examine the relationship between the sentiments extracted from the posts and student achievement. Before focusing on the sentiments, we are interested in finding out whether forum discussants would outperform other students in final grades. Table 5 gives the mean grades of forum discussants and other students as well as the *P*-value of Welch Two Sample *t*-test on the grades.

Table 5: Comparison of the grades of forum discussants and other students.

Measures	Values
Mean Grades of forum discussants	32.6
Mean grades of others	0.6
Welch Two Sample t-test P value	< 0.001

From Table 5 we learned that forum discussants did gain significantly higher grades compared with other students. Subsequently, using R 3.4.1 (Team, 2017), we regressed final grades on the sentiments and other behavioral variables in Table 4. Before inputting the data into the regression model, we transformed all the variables into [0, 1] using Max-Min method. The regression results are summarized in Table 6.

Table 6: Details of the trained regression model.

	Coefficients	*t* values	*P* values	
Intercept	0.10664	4.158	<0.001	***
Poscore	-0.14143	-1.185	0.2368	
Negscore	**0.23024**	**1.846**	**0.0655**	*
TLv	0.72383	9.055	<0.001	***
TLd	0.55621	5.368	<0.001	***
TFr	1.37113	5.035	<0.001	***
TFp	-0.32921	-1.034	0.3015	

(*P*⩽0.01: ***; 0.01<*P*⩽0.05: **; 0.05<*P*⩽0.1: *)

Table 6 indicates that *TLv*, *TLd* as well as *TFr* significantly correlate with *grade*, which is in line with existing studies (Yang et al., 2013; Zhuoxuan et al., 2015). With those variables controlled, *Poscore* is unrelated to *grade* while *Negscore* is significantly positively related to *grade*. One unit increase of *Negscore* leads to 0.23 unit increase of *grade* when other variables remain unchanged. Interestingly, the conclusion is opposite to extant research claiming that negative sentiments downgraded student achievements (Mega et al., 2014). To gain some insight on the positive influential effect of *Negscore* on *grade*, we first identified the negative words contributing most negative scores by ranking $TNeg_w$ of all the negative words in the forum posts, where $TNeg_w$ is defined as:

$$TNeg_w = -F_w \times s_w \tag{3}$$

. In Equation (3), F_w stands for the frequency of word *w* in all the posts. Some negative words with high ranks and their $TNeg_w$ values are listed in Table 7.

Table 7: Some negative words with high $TNeg_w$ values.

Negative words	$TNeg_w$
delete	20.3
wrong	19
worried	15
trouble	14.9
doubt	12

Then, we read the posts mentioning those negative words carefully, and found that most of the negative emotions are closely linked to assignments or quizzes instead of the learning attitude towards the course itself. Such posts suggest students' intention to figure out their problems with the help of other forum discussants. As a matter of fact, over 90% of students who have posted for help gained needed assistance, and this can be helpful in improving their final grades. We gave an example regarding student ***A***[3] in the course here. On November 15th, 2014, ***A*** initiated a post saying "Where goes wrong (the negative word) regarding my codes for swapping matrix rows? It is driving me crazy. Please give me some help." In the next five hours, ***A*** received two possible solutions from others. The first is about changing output order and the second is replacing the current swapping method. Finally, ***A*** solved his problem with the second solution. Over the course duration, ***A*** went through several rounds of similar experiences, and he gained the final grade of 88, achieving the "*distinction*" level.

4 CONCLUSION AND FUTURE WORK

This paper focused on investigating the relationship between student's emotional factors and achievement regarding a Chinese MOOC course. Positive and negative sentiments are extracted from the course forum with a lexicon-based method, and grades were regressed on the sentiments and other behavioral variables. The results showed that the positive sentiment was irrelevant to the grade, whereas the negative sentiment was significantly positively related to the grade. One plausible explanation for the conclusion is that the negative sentiment reflects students' achievement anxiety as they needed the help of other forum discussants to address their own problems. After receiving needed assistance, they would gain better grades with a higher probability. Based on the conclusion, some suggestions can be provided on adjusting course designs. First, incentives should be devised to encourage students to post their problems encountered during learning process actively, no matter how negative the problems appear. Second, incentives should also be devised to encourage forum discussants to help each other actively, as over 60% of the assistance to the students was given by TAs currently, which might take too much of their time and lead to inefficiency problem.

We have merely studied one MOOC course at present. In the future, we will analyze the data of more courses in other disciplines and find out whether the conclusion in this paper would still stand.

[3]For privacy purpose, we do not give details of the student.

REFERENCES

Chen, Q. (2007). Relationship between Emotional Intelligence, Achievement Motive and Academic Ahievments of University Students. *Journal of Higher Education Management, 1*(4), 78-81.

Chen, R., & Lazer, M. (2013). Sentiment analysis of twitter feeds for the prediction of stock market movement. *stanford. edu. Retrieved January, 25*, 2013.

Koller, D. (2012). MOOCs on the Move: How Coursera Is Disrupting the Traditional Classroom (text and video). *Knowledge@ Wharton Podcast. University of Pennsylvania, 7*.

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*(1), 121.

Pritchard, M. E., & Wilson, G. S. (2003). Using emotional and social factors to predict student success. *Journal of college student development, 44*(1), 18-28.

Ramesh, A., Goldwasser, D., Huang, B., Daumé III, H., & Getoor, L. (2013). *Modeling learner engagement in MOOCs using probabilistic soft logic.* Paper presented at the NIPS Workshop on Data Driven Education.

Team, R. C. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Villavicencio, F. T., & Bernardo, A. B. (2013). Positive academic emotions moderate the relationship between self‐regulation and academic achievement. *British Journal of Educational Psychology, 83*(2), 329-340.

Wang, S.-M., & Ku, L.-W. (2016). *ANTUSD: A Large Chinese Sentiment Dictionary.* Paper presented at the LREC.

Yang, D., Sinha, T., Adamson, D., & Rosé, C. P. (2013). *Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses.* Paper presented at the Proceedings of the 2013 NIPS Data-driven education workshop.

Zhao, S., Tong, Y., Liu, X., & Tan, S. (2016). *Correlating Twitter with the stock market through non-Gaussian SVAR.* Paper presented at the Advanced Computational Intelligence (ICACI), 2016 Eighth International Conference on.

Zhuoxuan, J., Yan, Z., & Xiaoming, L. (2015). Learning behavior analysis and prediction based on MOOC data. *Journal of Computer Research and Development, 52*(3), 614-628.

What's (not) in a Keystroke? Automatic Discovery of Students' Writing Processes Using Keystroke Logging

Rianne Conijn[1], Janneke van der Loo[2], Menno van Zaanen[1]
[1]Department of Cognitive Science and Artificial Intelligence
[2]Department of Cognition and Communication
Tilburg University
{m.a.conijn, j.m.vdrloo, m.m.vanzaanen}@tilburguniversity.edu

ABSTRACT: Teachers typically do not have access to students' writing processes, such as planning and revision, but only to final products. Students' writing processes can be analyzed by labor-intensive methods such as thinking aloud or by manually labelling behavior logs. This paper describes an approach to automatically extract writing processes from keystroke data. Keystroke data from 70 students writing an academic synthesis task are analyzed. A heuristic-based method is used to extract the planning and revision processes. In addition, Bayesian correlational analysis and *t*-tests are used to identify the relation between the extracted processes and students' self-reported writing style. The results show that the heuristic-based method can extract planning and revision features from keystrokes. However, no relation between the planning features and self-reported planning style and a limited relation between revision features and self-reported revision is found. Some anecdotal evidence is found that high revisers typed more revision characters than low revisors. To arrive at the fully automatic analysis of students' writing processes, future work should extract more keystroke features and evaluate their relation with the actual writing processes.

Keywords: Writing analytics, writing processes, writing strategies, keystroke analysis.

1 INTRODUCTION

Writing teachers often only have access to the final writing products constructed by the students, which does not include information about the actual writing processes. Two writing processes or strategies often used in writing research are planning and revision (Flower & Hayes, 1980). To improve writing instruction, it would be useful to have insight into these writing processes as well. Traditionally, students' writing processes were analyzed using thinking-aloud methods, self-report questionnaires, and retrospective interviews. Nowadays, with learning and writing becoming more digitalized, data about students' writing processes can be collected automatically. Keystroke logging is one tool which can be used to automatically collect students' typing behavior.

Keystroke logging in writing research has been used for a wide variety of aims. For example, keystroke logging has been used to predict essay score (Zhang, Hao, Li, & Deane, 2016), distinguish skilled versus less-skilled writers (Xu & Ding, 2014), to determine boredom and engagement (Allen et al., 2016), to assess mental ability (Van Waes, Leijten, Mariën, & Engelborghs, 2017), and to determine the tasks' cognitive load (Wallot & Grabowski, 2013). Yet, it is still considered difficult to extract higher-level writing processes from keystroke logs (Baaijen, Galbraith, & De Glopper, 2012; Leijten & Van Waes, 2013).

Some researchers tried to relate keystrokes to higher-level writing processes. Van Waes, Van Weijnen, and Leijten (2014) analyzed the relation between keystrokes and students' self-reported learning style when writing a bad news letter. No relation was found between the keystroke features (pauses between keys, characters produced) and learning style. Baaijen and colleagues (2012) did find a relation between keystroke features (timing of pauses, timing and place of revisions) and type of revisions. Using principal components analysis, five main components were derived: planned sentence production, within-sentence revision, revision of global structure, and (tentatively labeled) post draft revision and careful word choice. Lastly, Tillema, Van den Bergh, Rijlaarsdam, and Sanders (2011) related keystrokes to planning and revision behavior, with manual labels. In addition, they compared this behavior with self-reported planning and revision styles. High planners were found to read less often, were more likely to plan at the start than in the end, produced more, and revised more, compared to low planners. High revisers were found to read their own text less often, compared to low revisers.

In contrast to the studies above, we will automatically extract both planning and revision processes from keystroke data obtained during an academic synthesis task using a heuristic-based method. In addition, we try to relate the extracted writing processes to students' self-reported writing style.

2 METHOD

2.1 Participants

In this study, first year undergraduate communication and information sciences students from Tilburg University, who followed the course Academic Dutch were asked to complete an academic synthesis task in Microsoft Word. Demographics and self-reported writing styles were collected in the form of a pre-test. In total, 74 participants provided informed consent and participated in this study. The academic synthesis task is a mandatory task in the course. The task aims to practice writing an academic introduction. The participants were asked to read three short academic texts at home first. Thereafter, in the classroom, the participants got 30 minutes to write (the start of) an introduction in Dutch (their native language) based on these three academic texts. They were asked to type everything and to not make written notes. During this task, keystrokes were collected. After the task, the students were allowed to finish the task at home, before handing it in.

2.2 Writing style self-report

Students' self-reported writing style were collected with the Writing Style Questionnaire (Kieft et al., 2006; 2008). This questionnaire consists of 13 statements on planning, 12 statements on revision, and 12 filler statements. All questions were answered on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). This questionnaire provides two scores, a score on planning and revising style. Participants could score equal on both styles, or one of the styles could be dominant. The internal consistency was similar to that found by Kieft et al. (2006; 2008), with a Cronbach's alpha of .73 for the planning dimension and .69 for the revision dimension. The planning and revision scores were only moderately correlated (r = .39), indicating that the scores can be analyzed separately. The participants scored somewhat higher on revision (M = 3.5, *S.D.* = 0.48) compared to planning (M = 3.0, *S.D.* = 0.51). Median split was used to recode the planning and revision scores into binary variables, to analyze the differences in planning and revision between high and low planners and high and low revisers.

2.3 Keystroke data feature extraction

The keystrokes were collected with Inputlog (Leijten & Van Waes, 2013), which logs every key pressed and the times of the key press and key release. On average, the participants pressed 2967 keys (*S.D.* = 1076), which resulted in 2567 characters (*S.D.* = 1178) produced. The final document (after the 30 minutes) consisted on average of 1809 characters (*S.D.* = 895), indicating that a fair amount of revision took place. Planning and revision features were extracted from the keystrokes using a heuristic-based method. Here, rules are used to denote parts of the overall keystroke sequences as either 'planning' or 'revision'. On average, 2.9% of all the keystrokes were labeled as planning, and 16.4% of the keystrokes as revision.

The rules for labeling a sequence as *planning* included: the first phrases or non-complete sentences (sentences without a period) with at least 20 characters. Thus, typing a heading such as "inleiding opdracht 1" (*Dutch for: Introduction assignment 1*), would not be considered planning. Note, this only includes initial planning, not the planning in the middle of the writing processes when already some full sentences are produced. Based on these rules, four planning features were extracted: initial pause time (time until the first keystroke), number of plan characters, plan time, and plan character ratio (number of characters planned/total number of characters).

Rules for labeling a sequence as *revision* included all consecutive keystrokes where the next keystroke resulted in a lower document length, i.e., something was removed. Based on these rules, four revision features were extracted, which were similar to the planning features: the number of revisions, number of characters revised, revision time, and revision character ratio (number of characters revised/total number of characters).

Data from four participants were removed. One participant wrote in English instead of Dutch. In addition, three outliers (features more than three *S.D.* above the mean) were removed because these had a significant influence on the results. In total, data from 70 participants were left for analysis.

2.4 Analysis

The relation between the extracted planning and revision features with the self-reported writing style was analyzed using Pearson's correlation analysis, and evaluated with Bayes Factor, calculated in R with a Jeffreys-Zellner-Siow (JZS) prior set-up (Wetzels & Wagenmakers, 2012). The Bayes Factor (BF_{10}) quantifies the evidence in favor of one hypothesis, over an alternative hypothesis. The number indicates how much more (un)likely the data are to have occurred under the alternative hypothesis, compared to the null hypothesis. Next to correlational analysis, we analyzed whether high self-reported planners showed significantly more planning than low planners, and whether high self-reported revisers showed significantly more revisions compared to low revisers. Bayesian *t*-tests, implemented using the BEST package in R (Kruschke & Meredith, 2017), were used to compare the planning and revision features between the low/high planners and low/high revisers, respectively.

3 RESULTS

The correlational analysis showed that there is moderate evidence against a correlation between planning score and initial pause time and planning (r = -.09, BF_{10} = 0.13), the number of plan characters (r = .13, BF_{10} = 0.16), plan time (r = .07, BF_{10} = 0.10), and plan character ratio (r = .34, BF_{10} = 6.14). Here, a Bayes Factor of 0.13 indicates that the data are 1 / 0.13 = 7.7 times more likely to have occurred under the null hypothesis H_0 (no correlation) than under the alternative hypothesis H_1 (correlation). Thus, none of the planning features seem to be correlated with the self-reported planning score. For revision, a moderate evidence against a correlation was found between revision score and the number of revisions (r = .06, BF_{10} = 0.10), number of revision characters (r = .11, BF_{10} = 0.14), revision time (r = .08, BF_{10} = 0.12), and revision character ratio (r = .07, BF_{10} = 0.11).

In addition, Bayesian *t*-tests were conducted to analyze whether high self-reported planners or high revisers indeed showed more planning or revision compared to low planners or low revisers. No significant differences were found between the planning features for high/low planners (Table 1). All 95% highest density interval (HDI) included zero, thus the differences between the two means were not significantly different from zero. The low Bayesian Factors also support the evidence for the null model (no differences between the means). Likewise, all 95% HDIs for the reviser features included zero, indicating no significant differences between the revision features for high/low revisers (Table 2). However, the Bayes Factor does show some anecdotal evidence for a difference in revision characters. Further inspection indeed showed that 92% of the HDI was above zero, thus there is a 92% probability that the mean of the number of revision characters is higher for high revisers, compared to low revisers.

Table 1: Bayesian *t*-tests planning features

Feature	Overall M (S.D.)	High Planner M (S.D.)	Low planner M (S.D.)	95% HDI	BF_{10}
Initial pause time (s)	305 (105)	290 (87)	318 (118)	[-70.9, +25.4]	0.41

Number of plan characters	74 (116)	82 (119)	66 (115)	[-13.3, +17.3]	0.28
Plan time (s)	88 (151)	83 (156)	93 (149)	[-4.84, +11.4]	0.26
Plan character ratio (%)	3.5% (7.3%)	3.8% (6.7%)	3.2% (8.0%)	[-0.3%, +1.4%]	0.26

Table 2: Bayesian *t*-tests revision features

Feature	Overall M (S.D.)	High reviser M (S.D.)	Low reviser M (S.D.)	95% HDI	BF_{10}
Number of revisions	117 (55)	127 (67)	106 (44)	[-12.1, +46.0]	0.49
Number of revision characters	817 (653)	993 (873)	677 (358)	[-82.8, +460]	1.47
Revision time (s)	146 (64)	153 (69)	140 (59)	[-16.8, +42.4]	0.32
Revision character ratio (%)	30% (14%)	33% (15%)	28% (12%)	[-2.0%, +12%]	0.69

4 CONCLUSION

The current work described an approach to extract planning and revision processes from keystroke logs using a heuristic-based method. We showed that revision and planning processes can at least to some extent be extracted from keystrokes. In addition, properties from these processes were related to the self-reported planning and revision writing styles. In future work, we will explore ways to extract more (detailed) processes from keystrokes. For example, we will include planning in the middle of the writing task (we now only included initial planning) or different types of revision, such as surface and meaning revisions (Faigley & Witte, 1981). Yet, these features might be harder to accurately identify using a heuristic-based method.

The extracted keystroke features showed limited to no relation with the self-reported writing style. Only some evidence is found that high revisers use more revision characters, compared to low revisers. These findings are consisted with the findings Van Waes and colleagues (2014), who did not find a relation between self-reported learning style and keystroke features. However, Tillema et al. (2011) found a relation between self-reported writing style and keystroke features. Yet, in their study, the keystrokes were manually labeled with writing processes. This indicates that there is some relation between the self-reported writing style and actual planning and revision behavior. This would suggest that we are not yet extracting the right features from the keystrokes which represent planning and revision behavior. To evaluate whether the extracted features indeed relate to writing and revision processes, in future work, we will manually code the dataset to evaluate the extracted features. In addition, this labeled dataset can be used to automatically classify a given sequence of keystrokes. This paper showed the first steps towards the automatic discovery of writing processes using keystroke logging.

REFERENCES

Allen, L. K., Mills, C., Jacovina, M. E., Crossley, S., D'Mello, S., & McNamara, D. S. (2016). Investigating boredom and engagement during writing using multiple sources of information: the essay, the writer, and keystrokes. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 114–123). ACM.

Baaijen, V. M., Galbraith, D., & de Glopper, K. (2012). Keystroke Analysis: Reflections on Procedures and Measures. *Written Communication*, *29*(3), 246–277.

Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, *32*(4), 400–414.

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, *31*(1), 21–32.

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2006). Writing as a learning tool: Testing the role of students' writing strategies. *European Journal of Psychology of Education*, *21*(1), 17–34.

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2008). An aptitude–treatment interaction approach to writing-to-learn. *Learning and Instruction*, *18*(4), 379–390.

Kruschke, J. K. & Meredith, M. (2017). BEST: Bayesian Estimation Supersedes the t-Test. R package version 0.5.0. https://CRAN.R-project.org/package=BEST

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. *Written Communication*, *30*(3), 358–392.

Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2011). Relating self reports of writing behaviour and online task execution using a temporal model. *Metacognition and Learning*, *6*(3), 229–253.

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior*, *73*, 311–319.

Van Waes, L., van Weijen, D., & Leijten, M. (2014). Learning to write in an online writing center: The effect of learning styles on the writing process. *Computers & Education*, *73*, 60–71.

Wallot, S., & Grabowski, J. (2013). Typewriting Dynamics: What Distinguishes Simple From Complex Writing Tasks? *Ecological Psychology*, *25*(3), 267–280.

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064.

Xu, C., & Ding, Y. (2014). An exploratory study of pauses in computer-assisted EFL writing. *Language Learning & Technology*, *18*(3), 80–96.

Zhang, M., Hao, J., Li, C., & Deane, P. (2016). Classification of Writing Patterns Using Keystroke Logs. In L. A. van der Ark, D. M. Bolt, W.-C. Wang, J. A. Douglas, & M. Wiberg (Eds.), *Quantitative Psychology Research: The 80th Annual Meeting of the Psychometric Society, Beijing, 2015* (pp. 299–314). Cham: Springer International Publishing.

Beyond Identifying Areas for Improvement in Schools: Using the NILS Online Platform to Accelerate Improvement Work

Ouajdi Manai
The Carnegie Foundation for the Advancement of Teaching
manai@carnegiefoundation.org

Hiroyuki Yamada
The Carnegie Foundation for the Advancement of Teaching
yamada@carnegiefoundation.org

ABSTRACT: We confront a growing chasm between rising aspirations for our educational systems and what schools can routinely accomplish. Although educators at the classroom, school, and district levels are expending significant energy generating and testing promising interventions, we often observe the failure to scale up research-based knowledge across varied contexts. This interactive half-day workshop presents a way to move from trying to get better to getting good at getting better. We will introduce an improvement science approach that focuses on learning-by-doing to make progress toward a specific aim on a shared problem of practice by leveraging the power of networked communities. We will present how to apply the six core principles of improvement and organize improvement work through an online technology called NILS, emphasizing that (a) knowledge about the innovation itself and associated know-how around effective implementation flow through the interpersonal relationships between different actors; (b) attending to variation in performance and seeing the system that produces the current outcomes help us to identify areas for improvement. Utilizing NILS, participants will engage in structured activities and data exercises, learn how to identify areas for improvement from data, and create a driver diagram as a theory of practice improvement.

Keywords: Networked Improvement Community, Improvement Science, Social Learning, See the System, Systems Thinking, Scaling Up, Variation in Performance, Knowledge Dissemination.

1 BACKGROUND

We currently face a growing rift between rising expectations of what we want schools to achieve and what they can realistically accomplish. For instance, one of the main challenges in education is the failure to scale up research-based knowledge across varied contexts (Lewis, 2015). Bryk (2015) argues that we need an improvement paradigm that recognizes the complexity of educational work and the variability in educational outcomes that the current systems generate. Following this posit, over the past decade the Carnegie Foundation for the Advancement of Teaching has pioneered a fundamentally new vision for the research and development enterprise in education, seeking to join the discipline of improvement science with the powerful capacities of networks to foster innovation and social learning for education reform (Bryk, Gomez, Grunow, & LeMahieu, 2015).

Improvement work is organized around six core principles (Bryk et al., 2015): (a) make the work problem-specific and user centered (what specifically is the problem we are trying to solve?); (b) focus on variation in performance (what works, for whom, and under what set of conditions); (c) see the system that produces the current outcomes (*how local conditions shape work processes*); (d) embrace practical measurement (we cannot improve at scale what we cannot measure); (e) anchor practice improvement in disciplined inquiry (engage rapid cycles of PDSA [Plan, Do, Study, Act]); (f) accelerate improvements through networked communities (connect members of professional communities as the wisdom of crowds). Carnegie's approach to improvement is embodied in Networked Improvement Communities (NICs; Bryk et al., 2015). A NIC comprises a group of practitioners, administrators, researchers, and improvement specialists that works to improve a specific problem, shares a working theory of improvement embedded in systems thinking, and uses common measures and inquiry tools for learning whether the changes introduced are moving in the right direction.

NILS is the Networked Improvement Learning and Support online system developed by the Carnegie Foundation to accelerate the initiation and development of work in NICs. The impetus for building NILS emerged from listening to the needs and challenges addressed in various improvement communities including teachers, district leads, and state heads of education, where technology could be of great help in surmounting obstacles or catalyzing the improvement work. This platform is designed to align with the six core principles of improvement and follows the *SECI* model of promoting social, organizational learning and disseminating tacit and explicit knowledge (Nonaka & Takeuchi, 1995) for improvement in education by moving much of what we currently do face-to-face into a virtual learning environment. NICs are communities of practice and learning. Accordingly, NILS enables NIC members to learn improvement methods and culture within a system without the need for high-lift, in-person training. NICs initiate their work through seeing the system in the *Chartering* phase with in-site scaffolding for chartering activities. NICs then progress to system work in the *Improvement Testing* phase with a driver diagram, through which members test and record results for change ideas by running PDSA cycles. Ideas and individual learnings from PDSAs are then spread to the community for social learning. Social learning occurs through school-to-school, school-to-network, and network-to-network conversations among NIC members, which in turn enhance collaboration across the NIC both horizontally and vertically. As a change idea is tested across a variety of contexts, improvement ramps form and individual learnings converge as system knowledge. Members of a network hub curate knowledge gleaned from testing under varied contexts and share findings with the rest of the network, which prompts ideation for further changes. At its core, NILS attempts to address the question of how to derive knowledge from a NIC's data collection cycles: specifically, how does a system surface knowledge and wisdom to the right person at the right time? The platform aims to provide relevant data to testers based on their contexts and site interactions, and enhance connected learning for professional communities, thus enabling participating educators to take evidence-driven next steps towards achieving a collective aim.

At the LAK17 conference we introduced the initial version of NILS (Author, 2017) and received much interest from participants. We are now ready to show the enhanced version of NILS to participants at LAK18. Through this proposed workshop, participants will simulate an improvement work on NILS in a deliberate and systematic manner as if they were playing the part of members of a NIC. We will embed data exercises throughout the workshop, so that participants will learn how to identify areas for improvement and share their observations and learning with each other. In sum, our aim is to promote a more disciplined approach to improvement in schools by leveraging technology that supports continuous improvement.

2 ORGANIZATIONAL DETAILS

We propose a half-day, open workshop for applied researchers, evaluators, practitioners, and school leaders. Expected workshop activities are data exercises, discussion, and practice using NILS. We expect up to 40 participants, and plan on recruiting attendees via email and Twitter with the message, “Unleash the power of a Networked Improvement Community to coordinate your research efforts in a practical manner”. Required materials include the Internet, a laptop, a browser, as well as any educational problem that participants are trying to solve. The proposed agenda is presented below in Table 1.

Table 1: Proposed agenda

Session	Time	Content
1. Introduction	09:00 - 09:30	• Improvement Science • NIC life cycles
2. Launch the Simulation	09:30 - 10:00	• Simulation context & problem • Data conversation protocol • Forming a team
3. Understanding the problem	10:00 - 10:25	• Problem analysis
BREAK	10:25 - 10:45	• Sip tea/coffee (with snack)
4. Focusing Collective Efforts	10:45 - 11:10	• Consolidating learning • Formulating an aim statement
5. Change Ideas and Testing	11:10 - 11:45	• Developing change ideas • Formulating a driver diagram • Running a PDSA
6. Evidence in Improvement Science	11:45 - 12:10	• Multiple ramps of PDSAs • Assessing confidence to scale and

		spread
7. Unpacking NILS	12:10 - 12:20	• Current features and future roadmap
8. Closing remark	12:20 - 12:30	• Summarize key takeaways • Final reflection • Q & A

3 OBJECTIVE

Through this workshop, participants will be able to address (a) what it feels like to apply an improvement science approach to solving educational problems, (b) what skills and capacities will be required to do improvement work in educational context, and (c) how technology can be leveraged to organize and accelerate improvement work, and learn ways to identify areas for improvement and formalize a theory of improvement. We will go through those improvement principles mentioned above in a mode of "learning by doing" simulation and illustrate how each principle is manifested in the improvement work. Under each principle, we will introduce tools, examples, and data exercises to help participants illuminate variation and see the system. By the end of the workshop, participants will understand how to identify improvement priorities from data and create a driver diagram as a representation of their theory of improvement. We will introduce NILS as an online tool designed to support this work process. We will provide an overview of these learning outcomes through series of tweets using the #improve hashtag. In addition, we will send an email reminder with a detailed agenda and logistics for the workshop.

REFERENCES

Author. (2017, March).

Bryk, A. S. (2015). Accelerating How We Learn to Improve. *Educational Researcher, 44*(9), 467-477. https://doi.org/10.3102/0013189X15621543

Bryk A. S., Gomez L. M., Grunow A., LeMahieu P. G. (2015). *Learning to improve: How America's schools can get better at getting better.* Cambridge, MA: Harvard Education Press.

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54-61. https://doi.org/10.3102/0013189X15570388

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation.* New York, NY: Oxford University Press.

Now You See It: Optimizing Student Performance with Learning Environment Modeling and Insight Reports

Liz Crowell
University of Central Oklahoma
ecrowell1@uco.edu

Bucky Dodd
Institute for Learning Environment Design
bdedd1@uco.edu

ABSTRACT: In this hands-on workshop, participants will explore Learning Environment Modeling™ (LEM); and gather, analyze, and report data in a meaningful way. Learning Environment Modeling™ is a way of visually representing a learning environment so it may be shared and easily understood regardless of area of expertise. Additionally, quantitative and qualitative data is gathered from students, faculty, and university systems, followed by analysis in order identify where and how students may not be successful and can improve. Participants will be able to build LEM patterns, and generate, analyze, and present information to support data-informed decisions.

Keywords: Learning Environment Modeling, LEM, Student Success, Satisfaction, Design

1 INTRODUCTION AND BACKGROUND

Designing learning environments and teaching are complicated, ever-changing processes. Communicating the design and facilitation is even more so, with usually a syllabus serving as the only, unchanging guide. Learning Environment Modeling™ offers a shared language that communicates the what, why, and how of a project that can be scaled to any level from assignment to course to university program (and beyond!). When LEM is used as a diagnostic tool and combined with quantitative and qualitative data from facilitators and students, the resulting Insight Report identifies problem points and empowers faculty to optimize student performance, in real-time and for the next group of students.

1.1 Learning Environment Modeling™

LEM finds its base in the Learning Environment Modeling Language or LEML. LEML features five symbols and four containers that allow faculty to collaborate regardless of subject matter (see Figure 1). The blueprints created with LEM allow educators and colleagues to diagnose, improve, and present learning activities. Additionally, LEM blueprints can be built in analog or digital form, increasing access and sharing capabilities, as seen in Figure 2.

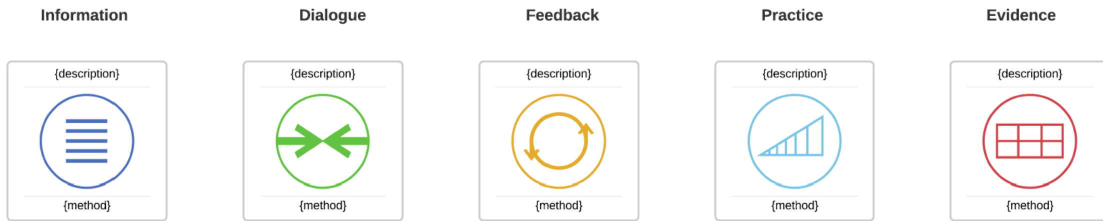


Figure 1: The five building blocks of Learning Environment Modeling Language

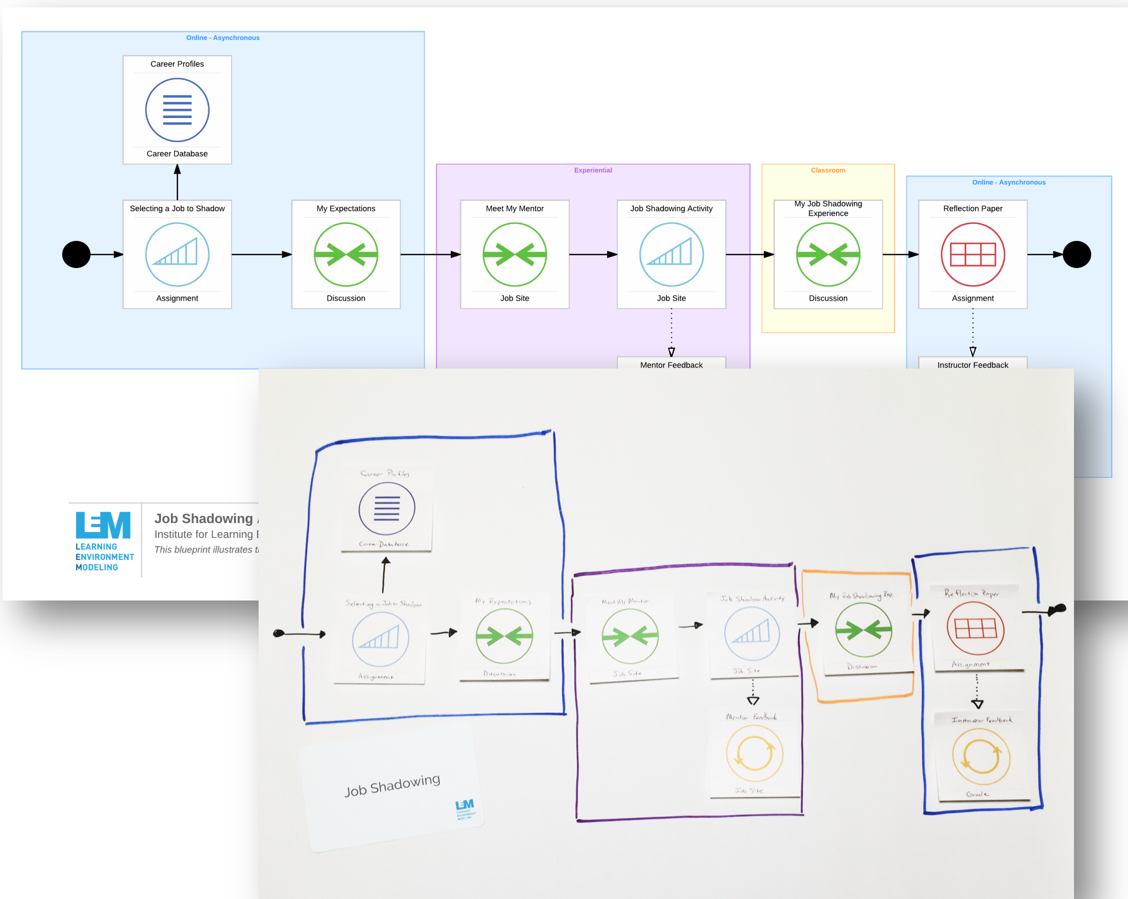


Figure 2: Learning Environment Modeling patterns represented in analog and digital formats

1.2 Data Sources and Insights

The patterns themselves are full of data points, as a whole and as individual units. Additional data sources include outside quality evaluation, student success, student satisfaction, and faculty satisfaction. Since all of these sources come from university systems, learning management systems, and third party tools, and contain many quantitative and qualitative data points; the key is discerning what is relevant and what is not. While the process of identifying, analyzing, and organizing the relevant data points to determine successful and not-so-successful elements is challenging, the results become very clear – to everyone. Instead of telling designers or facilitators what to do, LEM and the Insight Report present environment-specific information and recommendations in an objective, easy-to-understand way that gives the user full ownership (see Figure 3). Piloting begins in October 2017 with two groups completing by February 2018.

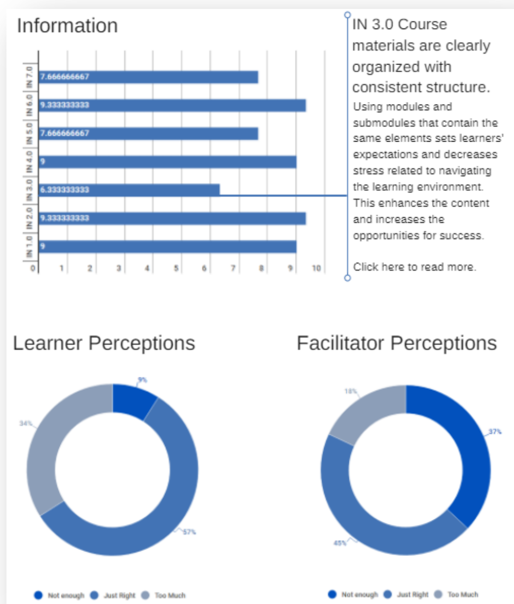


Figure 3: Page from an Insight Report indicating lower success with a specific point in a course.

2 ORGANIZATIONAL DETAILS OF PROPOSED EVENT

2.1 Type of event, proposed schedule, and duration:

This session is proposed as a full-day, hands-on workshop in the Understanding Learning & Teaching, Modeling track. There will be an introduction to LEM, pattern creation, and data sourcing in the morning followed by analysis and reporting in the afternoon.

2.2 Type of participation:

Participation is open.

2.3 Workshop activities:

This session will contain discussions, hands-on building of patterns in either analog or digital as well as the collection and meaningful reporting of data.

2.4 Expected participant numbers and planned dissemination activities to recruit attendees:

This workshop can scale from a few participants to a large group of 50 or more people. Social media methods will be used to recruit attendees and describe the benefits of the workshop. We will also use email nurturing campaigns to recruit attendees.

2.5 Required Equipment:

Projector and screen. All other materials will be provided.

3 OBJECTIVES

- Describe Learning Environment Modeling
- Identify components of the Learning Environment Modeling Language
- Build a pattern with Learning Environment Modeling using support materials
- Connect and analyze the pattern with user-provided, quantitative and qualitative data
- Analyze learning performance and satisfaction as well as design potential solutions
- Present tailored recommendations to optimize the pattern in an easy-to-understand report

3.1 Dissemination of outcomes:

All participants will receive Learning Environment Modeling™ guides and support materials that can be disseminated in a variety of media, based on organizer or participant preference.

REFERENCES

Institute for Learning Environment Design. (2017). *The Learning Designer's Guide to LEM.* Oklahoma City, OK. Institute for Learning Environment Design

Dodd, B., Gillmore, J. (2015). *Learning Environment Modeling: Redefining Learning Environment Design.* Oklahoma City, OK. Institute for Learning Environment Design

Launching a Career in Learning Analytics

Danielle Hagood
University of California Davis, USA
dhagood@ucdavis.edu

Robert Bodily
Birgham Young University, USA
bodilyrobert@gmail.com

ABSTRACT: This half-day workshop addresses substantial interests graduate students expressed at LAK17. We aim to provide both an introduction to the learning analytics field and networking opportunities to connect participants with other scholars—promoting future collaboration. This workshop is an entry-point targeting graduate students, particularly those at institutions without learning analytics mentorship. Specific topics include an overview of research areas in learning analytics, exposure to domain specific social-science and computer-science skills, and examples of the "pipeline" of doing learning analytics projects. Workshop time will be split between presentations and structured networking and collaboration, with the latter promoting lasting connections among participants. We discuss how these community building efforts will be sustained beyond this workshop through a SoLAR SIG and report of resources.

Keywords: graduate students, professional development, interdisciplinary, communities of practice, learning analytics

1 WORKSHOP BACKGROUND

The impetus for this workshop began at LAK17, when over 20 graduate student and post-doc scholars met to discuss the need for training and collaborative mentoring in the field of learning analytics. Primary concerns centered around two issues. First, many pre-career scholars described training at institutions with advisors that have little to no expertise in learning analytics. Often, a student's initial interest in the learning analytics body of research is stymied by not knowing where to start. Second, that discussion emphasized a desire for more connection with other scholars interested in learning analytics work. By strengthening a network of young scholars, we believe the future learning analytics research will grow more vibrant. We—like scholars expecting an increasingly interdisciplinary field (see Dawson, Gašević, Siemens, & Joksimovic, 2014)—see new PhDs as bringing diversity of methods and ideas to the discipline. To sustain network building resulting from this workshop, a Society of Learning Analytics Research (SoLAR) Graduate Student Special Interest Group (SIG) has been formed. The primary purpose this Graduate Student SIG achieve with this workshop, is to provide a jump start to graduate students that do not have learning analytics support at their institution.

The material for this workshop build on peer information sharing in the community-of-practice model (Wenger, 2011). The organizers took initiative starting the SoLAR Graduate Student SIG because the interests and struggles described at LAK17 resonated with our own experiences. Because of our experiences, we curated information we have found useful in our own processes of discovering learning analytics alone. To validate and supplement our ideas—and to prevent blind leading the blind—we are conducting informal interviews with experienced learning analytics scholars, in preparation for this workshop.

Finally, the existing SoLAR doctoral consortium offers an excellent platform for graduate students to receive feedback on their work, once they start research in the field of learning analytics. However, we think newer members of our community also have questions about how to get started, especially graduate students with little to no access to learning analytics scholars. This workshop serves as an intense, half-day introduction to learning analytics, and we hope, a community building exercise for young scholars.

2 ORGANISATIONAL DETAILS

2.1 Type of Event

This proposed event is a half-day workshop. The coordinators will balance interactive activities with information presentation. Workshop attendees will collectively participate in several un-workshop style activities.

2.2 Schedule and Activities

2.2.1 Welcome and Survey—30 minutes

To begin, we want participants to introduce themselves, since a goal of this workshop is to build connections. The participants will then complete a reflective survey including general questions they bring to the workshop, research interests, areas of expertise, training needs, and future career plans. Results from this survey will be discussed later in the workshop.

2.2.2 What Does a "Learning Analyst" Need to Know?—40 minutes

Next, we will present a brief 20 minute talk summarizing the broad topics subsumed within learning analytics. For this talk, we are drawing on resources like *The Handbook of Learning Analytics* (Lang, Siemens, Wise, & Gašević, 2017) and InfoHub to identify themes in the field and create a "short list" of preeminent articles relevant to each area for dissemination. Based on that list and the "What Does a Learning Analysts Need to Know" presentation, we will then facilitate discussion around what areas of work are developing in learning analytics. With this discussion, the underlying emphasis is that in learning analytics, no one scholar does it all. Leaving participants with the understanding that interdisciplinary collaboration is essential. We hope that participants will relate their work to an area of inquiry in the field.

2.2.3 Learning Science and Computer Science—60 minutes

From a broad presentation and discussion on areas of work in the field, we transition to focus on particular types of skills that should be fostered in graduate school. In particular, we have found that many graduate programs emphasize either social-science and learning-science skills or computational skills. Thus, we anticipate splitting participants into two groups based on which of these skillsets they feel most comfortable. Then, we have two different brief talks and collaborative activities planned to expose participants to the "other side's" methodology and epistemology. We start with the methodological orientation and perspectives, discuss particular analytic skills, and finally conclude with some advice on how to collaborate across disciplines based on interdisciplinary work research (Klein, 2010; Mansilla & Duraising, 2007).

2.2.4 Example Learning Analytic Research Pipeline—60 minutes

Often, starting a research project is daunting and exciting as a graduate student. But when entering a new field, the presenters both felt unguided and often did not know what research process looked like behind the articles that we read. Drawing on the struggles in our own dissertation work and especially on interviews with established authors, we will present a general model workflow for several learning analytics projects. We will present several JLA articles and summarize the technical aspects underlying that work, according to the authors. From this, we hope to highlight issues particular to learning analytics work, such as participant privacy and managing large data. A secondary goal is to evidence how diverse research in learning analytics can be. This section will conclude with small group discussion of participants' work and peer feedback on research design (participants will be asked to bring a one-page description of their current research interests).

2.2.5 Your Future in Learning Analytics—50 minutes

The last section of the workshop will briefly return to the survey participants completed at the beginning. We will look over the themes, both strengths and struggles, shared among participants. From this point, we will have a short 15-minute presentation describing a "day-in-the-life" of both academic and industry individuals working in learning analytics, based on informal interviews conducted by the coordinators. We also have a handout of curriculum vitae (CV) and dissertation "to do's" based on expert feedback. After this, we have will engage workshop participants in discussing their future career interests. The objective is to have individuals identify other participants with shared interests to build network connections. Finally, we will share information about how to be involved with SoLAR, including a the Graduate Student SIG (we will be sure to mention the happy hour the SIG is planning at LAK18).

2.3 Recruitment and Dissemination

This event will be promoted through the SoLAR Graduate Student SIG email list and the learning analytics slack channel. We believe this workshop will hold special interest to graduate students and new scholars. Our intended participant size is 10 to 30 attendees. To further motivate students to attend our workshop, the SoLAR Graduate Student SIG can offer five micro-scholarships to workshop attendees. These micro-scholarships ($150 USD) will offer some financial support to new graduate students, with a particular focus on first time graduate student LAK attendees that do not have easy access to learning analytics support through their adviser, department, or university.

2.4 Equipment

No special equipment will be needed beyond audio / visual presentation equipment. Printed workshop materials for activities will be provided by the coordinators.

3 INTENDED OUTCOMES

An intangible outcome of this workshop will be welcoming and connecting young scholars interested in learning analytics. Concrete measures of this goal include recruiting new members to the SoLAR Graduate Student SIG, recruiting attendees to future LASI events, and forming research connections leading to presentations at future LAK conferences or even JLA publications. We expect a number of individual outcomes for participants. Though the workshop will not teach a specific analytic tool or method, we hope participants leave understanding the scope of learning analytics research, the various types of skills utilized in this interdisciplinary work, and how to find information and mentorship. Finally, at a more affective level, we hope that the individual outcomes for participants include increased confidence in participating in learning analytics research, connections to other scholars with similar research interests, and a sense of belonging in the SoLAR community. Further, the coordinators for this workshop wish to organize the information shared in a brief report. This resource can be disseminated through the SoLAR Graduate Student SIG network. Our preference is to create an open document that will be alive for revisions over time. This report will also be reflective, including survey responses from participants (and other Graduate Student SIG respondents) regarding research interests, areas of expertise, training needs, and future career plans. We also plan to compile what we learn from expert interviews and participant surveys into a research paper, investigating how to best support graduate students in the field of learning analytics.

REFERENCES

Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014, March). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 231-240). ACM.

Klein, J. T. (2010). *A taxonomy of interdisciplinarity.* The Oxford handbook of interdisciplinarity, 15.

Lang, C., Siemens, G., Wise, A., & Gašević, D. (2017). *The handbook of Learning Analytics*. Retrieved from: https://solaresearch.org/hla-17/

Mansilla, V. B., & Duraising, E. D. (2007). Targeted assessment of students' interdisciplinary work: An empirically grounded framework proposed. *The Journal of Higher Education, 78*(2), 215-237.

Wenger, E. (2011). Communities of practice: A brief introduction. National Science Foundation.

Sharing and Reusing Data and Analytic Methods with LearnSphere

John Stamper
Carnegie Mellon University
john@stamper.org

Kenneth Koedinger
Carnegie Mellon University
koedinger@cmu.edu

Carolyn Rose
Carnegie Mellon University
cprose@cmu.edu

Philip Pavlik
University of Memphis
ppavlik@memphis.edul

ABSTRACT: This workshop will explore LearnSphere, an NSF-funded, community-based repository that facilitates sharing of educational data and analytic methods. The workshop organizers will discuss the unique research benefits that LearnSphere affords. In particular, we will focus on Tigris, a workflow tool within LearnSphere that helps researchers share analytic methods and computational models. Authors of accepted workshop papers will integrate their analytic methods or models into LearnSphere's Tigris in advance of the workshop, and these methods will be made accessible to all workshop attendees. We will learn about these different analytic methods during the workshop and spend hands-on time applying them to a variety of educational datasets available in LearnSphere's DataShop. Finally, we will discuss the bottlenecks that remain, and brainstorm potential solutions, in openly sharing analytic methods through a central infrastructure like LearnSphere. Our ultimate goal is to create the building blocks to allow groups of researchers to integrate their data with other researchers in order to advance the learning sciences as harnessing and sharing big data has done for other fields.

Keywords: Learning metrics; data storage and sharing; data-informed learning theories; modeling; data-informed efforts; scalability.

1 BACKGROUND

The use of data to improve student learning has become more effective as student learning activities and student progress through educational technologies are increasingly being tracked and stored. There is a large variety in the kinds, density, and volume of such data and to the analytic and adaptive learning methods that take advantage of it. Data can range from simple (e.g., clicks on menu items or structured symbolic expressions) to complex and harder-to-interpret (e.g., free-form essays, discussion board dialogues, or affect sensor information). Another dimension of variation is the time scale in which observations of student behavior occur: click actions are observed within seconds in fluency-oriented math games or in vocabulary practice, problem-solving steps are observed every 20 seconds or so in modeling tool interfaces (e.g., spreadsheets, graphers, computer algebra) in intelligent tutoring systems for math and science, answers to comprehension-monitoring questions are given and learning resource choices are made every 15 minutes or so in massive open online courses (MOOCs), lesson completion is observed across days in learning management systems, chapter/unit test results are collected after weeks, end-of-course completion and exam scores are collected after many months, degree completion occurs across years, and long-term human goals like landing a job and achieving a good income occur across lifetimes. Different paradigms of data-driven education research differ both in the types of data they tend to use and in the time scale in which that data is collected. In fact, relative isolation within disciplinary silos is arguably fostered and fed by differences in the types and time scale of data used (cf., Koedinger et al., 2012).

Thus, there is a broad need for an overarching data infrastructure to not only support sharing and use within the student data (e.g., clickstream, MOOC, discourse, affect) but to also support investigations that bridge across them. This will enable the research community to understand how and when long-term learning outcomes emerge as a causal consequence of real-time student interactions within the complex set of instructional options available (cf., Koedinger et al., 2010). Such an infrastructure will support novel, transformative, and multidisciplinary approaches to the use of data to create actionable knowledge to improve learning environments for STEM and other areas in the medium term and will revolutionize learning in the longer term.

LearnSphere transforms scientific discovery and innovation in education through a scalable data infrastructure designed to enable educators, learning scientists, and researchers to easily collaborate over shared data using the latest tools and technologies. LearnSphere.org provides a hub that integrates across existing data silos implemented at different universities, including educational technology “click stream” data in CMU’s DataShop (Stamper et al., 2011), massive online course data in Stanford’s DataStage and analytics in MIT’s MOOCdb (Veeramachaneni et al., 2014), and educational language and discourse data in CMU’s new DiscourseDB (Jo et al., 2016). LearnSphere integrates these DIBBs in two key ways: 1) with a web-based portal that points to these and other learning analytic resources and 2) with a web-based workflow authoring and sharing tool called Tigris. A major goal is to make it easier for researchers, course developers, and instructors to engage in learning analytics and educational data mining without programming skills.

This workshop builds off a successful LAK 2017 Tutorial, and workshop at AIED/EDM 2017. We hope that this year we will be able to attract attendees that have been exposed to LearnSphere from these past events, although we will have some tutorial activities included for new attendees as well.

2 ORGANIZATION

2.1 Type of Event

Workshop

2.2 Proposed Schedule and Duration

Table 1: Proposed Full Day Schedule.

Time	Item
8:30	Introductions
9:00	LearnSphere overview & high-level discussion I
10:00	Coffee Break
10:30	Tigris workflow tool (Lecture & Demos)
11:15	Hands-on I: Build custom analysis workflows using existing Tigris components

12:30	Lunch Break
1:30p	5-10 minute participant talks about proposed or created workflows
2:30p	Coffee Break
2:45p	Hands-on II: 2 breakout sessions (upload your own data; create workflow components)
4:15p	High-level discussion II
4:45p	Closing

2.3 Type of Participation

Mixed participation will be through submission of reviewed abstracts, invited guests, and open registration. For participants who have accepted abstracts or are invited by the workshop committee, we have allocated approximately $20,000 from our grant funding to cover registration and travel costs.

2.4 Activities

Activities will include presentations from workshop organizers, invited guests, and short presentations from accepted abstract presenters. Hands on sessions will include demos and group work towards implementing analytics.

2.5 Expected Numbers

We expect 15-20 participants based on previous workshops.

2.6 Activities to Recruit Attendees

We will create a website to announce the workshop and method of submitting abstracts. The Learning Analytics, Educational Data Mining, and LearnLab mailing lists will be used to direct potential attendees to the workshop website. In addition, we will invite a number of invited guests. Both accepted submissions and invited guests will have the chance to receive funding to attend.

2.7 Required Equipment

Projector and screen will be required by organizers. Attendees will need to bring laptops and will need adequate internet connectivity.

3 OBJECTIVES AND OUTCOMES

Broadly, this workshop offers those in the Learning Analytics community an exposure to LearnSphere as a community-based infrastructure for educational data and analysis tools. In opening lectures, the organizers will discuss the way LearnSphere connects data silos across universities and its unique capabilities for sharing data, models, analysis workflows, and visualizations while maintaining confidentiality.

More specifically, we propose to focus on attracting, integrating, and discussing researcher contributions to Tigris, the web-based workflow authoring and sharing tool. Workshop submissions in the form of abstracts will involve a brief description of an analysis pipeline relevant to modeling educational data as well as accompanying code. Prior to the workshop itself, the organizers will coordinate with authors of accepted submissions to integrate their code into Tigris. A significant portion of the workshop will be dedicated to hands-on exploration of custom workflows and workflow modules within Tigris. Authors of accepted submissions will present their analysis pipelines, and everyone attending the workshop will be able to access those analysis pipelines within Tigris to a variety of freely available educational datasets available from LearnSphere. The goal is to generate -- for each workflow component contribution in the workshop -- a publishable workshop paper that describes the outcomes of openly sharing the analysis with the research community.

Finally, workshop attendees will discuss bottlenecks that remain toward our goal of a unified repository. We will also brainstorm possible solutions. Our goal is to create the building blocks to allow groups of researchers to integrate their data with other researchers we can advance the learning sciences as harnessing and sharing big data has done for other fields.

REFERENCES

Jo, Y., Tomar, G., Ferschke, O., Rosé, C. P., & Gašević, D. (2016, April). Pipeline for expediting learning analytics and student support from data in social learning. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 542-543). ACM.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2010). The knowledge-learning-instruction (KLI) framework: Toward bridging the science-practice chasm to enhance robust student learning. Cognitive Science.Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2010). The knowledge-learning-instruction (KLI) framework: Toward bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*.

Stamper, J., Koedinger, K.R., Baker, R., Skogsholm, A., Leber, B., Demi, S., Yu, S., Spencer, D. (2011) Managing the Educational Dataset Lifecycle with DataShop*. In Kay, J., Bull, S. and Biswas, G. eds. Proceeding of the 15th International Conference on Artificial Intelligence in Education (AIED2011).*

Veeramachaneni, K., Halawa, S., Dernoncourt, F., O'Reilly, U. M., Taylor, C., & Do, C. (2014). Moocdb: Developing standards and systems to support MOOC data science. *arXiv preprint arXiv:1406.2015.*Gagné, M., Forest, J., Vansteenkiste, M., Crevier-Braud, L., van den Broeck, A., Aspeli, A. K., . . . Westbye, C. (2015). The Multidimensional Work Motivation Scale: Validation evidence in seven languages and nine countries. *European Journal of Work and Organizational Psychology, 24*(2), 178-196.

2nd Annual Workshop of the Methodology in Learning Analytics Bloc (LAKMLA18)

Yoav Bergner
New York University, USA
Yoav.bergner@nyu.edu

Charles Lang
Columbia University, USA
Cl3584@tc.columbia.edu

Geraldine Gray
Institute of Technology
Blanchardstown, Ireland
Geraldine.gray@itb.ie

ABSTRACT: Learning analytics is an interdisciplinary and inclusive field, a fact which makes the establishment of methodological norms both challenging and important. Building on the success of the LAK17 workshop on methodology this community-building workshop intends to convene methodology-focused researchers to discuss new and established approaches, comment on the state of current practice, author pedagogical manuscripts, and co-develop guidelines to help move the field forward with quality and rigor.

Keywords: Models, Methodology, Measurement, Statistics, Evaluation

1 WORKSHOP BACKGROUND

Learning analytics is an interdisciplinary and inclusive field that brings together educational technologists, psychologists, data scientists, learning scientists, substantive experts in various domains, and measurement specialists (Siemens and Gašević, 2012). For all of the strength that comes from such diversity, there are also potential pitfalls when it comes to establishing norms for methodological work. For example, Clow (2013) described learning analytics as, "a 'jackdaw' field of enquiry, picking up 'shiny' techniques, tools and methodologies… This eclectic approach is both a strength and a weakness: it facilitates rapid development and the ability to build on established practice and findings, but it—to date—lacks a coherent, articulated epistemology of its own." (p. 686).

In the years since this observation, the learning analytics community has grown rapidly, and the number of shiny techniques has grown as well. Looking just at recent proceedings of the International Conference on Learning Analytics and Knowledge (LAK), the variety is staggering. Methods range from descriptive statistics to correlation analyses, classification, clustering, regression, (M)AN(C)OVA, structural equation modeling, item response theory, hidden Markov models, time-series analysis, latent semantic analysis, social network analysis, and the list goes on. It is understandable and even expected that reviewers and readers of learning analytics manuscripts are unlikely to be expert evaluators of the methodological rigor in all of these cases. There is a naturally occurring process of specialization in any academic field. However, if growth of adoption outpaces systematic specialization then there is a risk that methodological errors will proliferate and that quality of community products will suffer.

To make matters even more complex, a number of recent papers have emphasized the sensitivity of quantitative analyses to data collection and variable operationalization choices, for example with regard to effects of selection bias (Brooks, Chavez, Tritz & Teasley, 2013), results of time-on-task analyses (Kovanović et al., 2015), studies of discussion forum usage (Bergner, Kerr, & Pritchard, 2015) and evaluation of student models (Pelánek, Rihák, & Papoušek, 2016). In addition, learning analytics models often incorporate a selection of proxy variables as indicators of latent constructs, such as learning and engagement. What proxy variables actually measure is less clear. For example, measures of engagement may be influenced by instructional conditions (Gašević, Dawson, & Siemens, 2015), adding ambiguity, and a lack of consistency, to our interpretation of models of learning.

In short, methodological concerns can arise from a range of practices including but not limited to selecting inappropriate methods, misusing methods, inadequate model evaluation or model comparison, sensitivity to operationalization, and over-reliance on proxy variables (Bergner, 2017). As the learning analytics community matures, it is particularly important to establish standards for good practice and to educate new students in accordance with these standards. Clear methodological guidelines increase the quality of work and facilitate communication not only within the community but also with practitioners in other research communities, where norms may be clearer. This is a challenging problem in large part because of the aforementioned diversity of approaches. The present workshop seeks to build a community of researchers with an interest in methodology and its rigorous application and development to the field of learning analytics.

There have been several previous LAK workshops and tutorials that have focused on specific methodologies—a limited set of examples includes the tutorials for classification and clustering using Weka (2014, 2016), special topics in discourse analysis (2013-2014) and writing analytics (2016-2017), assessment design (2016-2017), and temporal analysis (2012-2016)—but not on cross- cutting methodological issues such as developing methodological frameworks within learning analytics, framing and prioritizing methodological issues for the community, and providing resources to move the field forward.

1.1 Building on the LAK17 Workshop

The LAK17 Methodology Workshop received substantial interest from a variety of LAK participants, from seasoned computer scientists to people who were entering the field of learning analytics for the first time. The event served to seed a community that was interested in having both high level discussions of what methodology means in learning analytics and specific methodological issues that can arise in both quantitative and qualitative investigations. Arising directly from this workshop, the Journal of Learning Analytics current call for papers for a special section on methodological choices invites manuscripts on both of these topics. We plan to continue to build this community at LAK18 with an eye to segmenting the interest group into specific project-based groupings that may leverage the expertise present to generate impactful products. The LAK17 event began this process by defining possible projects, such as "cheat sheets" for relevant methods and the publication of methodological problems specific to the field, and we would like the opportunity to follow up on the progress made towards these aims as a larger group and determine ways that we can be more useful to the field as a whole.

1.2 Relevance to the Theme

We plan to incorporate the LAK18 theme into the workshop in various ways. Specifically, the discussion of user centered design as a methodology in itself and how and whether it is possible to incorporate it into all learning analytics work. A focus on methodological rigor also supports evidence based learning analytics practice.

2 ORGANISATIONAL DETAILS

The proposal is for a half day, open workshop covering introductions (10 mins), poster presentations (30 mins) and paper presentations (90 mins) interleaved by two group discussions (30 mins each). Including breaks, the session will last 4 hours in total. A call for papers and posters will be disseminated through relevant listservs, our network of contacts that have expressed an interest in methodology in learning analytics, and a workshop website. The expected participant number is approximately twenty.

3 WORKSHOP OBJECTIVES

Solicit Contributions Focused on Methodological Issues in Learning Analytics

The first objective of the present workshop, as before, is to solicit new substantive contributions specializing on methodology. In the first year of the workshop, we anticipated that contributions would fall roughly into the following categories: papers presenting new methods or adaptation/modification of methods; position papers which take a critical look at methodological practice in the community; and pedagogical/instructional papers oriented at students or researchers who are new to the field or developing an interest in a particular methodology.

Several participants in last year's workshop confirmed a need for products which provide compact, substantive guidance on methodological issues. We thus propose a new category of poster submissions for the LAK18 workshop, Methodology Guideline Posters (MGP).

Methodology Guidelines Posters

Related to the position and instructional papers that may be presented during the workshop, a second objective of convening will be to cooperatively develop community guidelines regarding the uses of various methods including data acquisition, data analysis and evaluation of results in conference and journal publications.

The idea behind Methodology Guideline Posters (MGPs) is that they should be infographic representations of decision flows in learning analytics methodology, working backwards from the ultimate goals. An MGP will emphasize how operational decisions are guided not only by the goals but also by the types and properties of available data and by problems of statistical inference. We do not imagine that MGPs will be instructional with regard to how to carry out analyses but rather will rather point the reader to appropriate references. The emphasis of MGPs will be interrogating the methodological choices. As such they should describe alternate case scenarios, explain pitfalls, and suggest options for sensitivity and goodness-of-fit tests.

Provide Expertise for Review Panels

A third objective of the workshop is to take responsibility for maintaining a database of methodology experts who are active in the learning analytics community. The expert listing is by no means intended to be exclusionary or to promote certain researchers over others but rather to help community members and editorial committees find methodology experts who are willing to consult and/or review relevant work.

Community Building

Last but not least, an objective of the workshop is to provide a meeting place for researchers who take a special interest in methodological issues. We anticipate that a concentrated meeting will promote continuing collaboration on this important topic.

REFERENCES

Bergner, Y. (2017) Measurement and its uses learning analytics. In Lang, C., Siemens, G., Wise, A., & Gašević, D. (Eds.), *Handbook of Learning Analytics* (pp. 34-48). Society for Learning Analytics Research.

Bergner, Y., Kerr, D., and Pritchard, D. E. (2015) Methodological Challenges in the Analysis of MOOC Data for Exploring the Relationship between Discussion Forum Views and Learning Outcomes. Proceedings of 8th International Conference on Educational Data Mining.

Brooks, C., Chavez, O., Tritz, J., and Teasley, S. (2015) Reducing selection bias in quasi-experimental educational studies. *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp 295–299), ACM.

Clow, D. (2013) An overview of learning analytics. *Teaching in Higher Education* 18(6), 683–695. Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R.S.J.D., and Hatala, M. (2015) Penetrating the black box of time-on-task estimation. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* - LAK '15 (pp. 184-193), ACM.

Gašević, D., Dawson, S., and Siemens, G. (2015) Let's not forget: Learning analytics are about learning. *TechTrends* 59(1), 64-71.

Pelánek, R., Rihák, J., and Papoušek, J. (2016) Impact of data collection on interpretation and evaluation of student models. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, (pp 40–47), ACM.

Siemens, G. and Gašević, D. (2012) Guest Editorial-Learning and Knowledge Analytics. *Educational Technology & Society* 15(3), 1–2.